UPDATED JULY 2023

AN UPDATED WEIGHTING STRATEGY FOR THE MONITORING THE FUTURE PANEL STUDY

MONITORING THE FUTURE OCCASIONAL PAPER NO. 98

Sponsored by the National Institute on Drug Abuse at the National Institutes of Health

monitoring the future

a continuing study of American youth

Megan E. Patrick
Yvonne M. Terry-McElrath
Patricia Berglund
Yuk C. Pang
Steven G. Heeringa
Yajuan Si

An Updated Weighting Strategy for the Monitoring the Future Panel Study

Megan E. Patrick, PhD[1,*]
Yvonne M. Terry-McElrath, MSA[1]
Patricia Berglund, MBA[1]
Yuk C. Pang, PhD[1]
Steven G. Heeringa, PhD[1]
Yajuan Si, PhD[1]

[1]Institute for Social Research, University of Michigan, Ann Arbor, MI
*Corresponding author: Megan E. Patrick, PhD., University of Michigan, Institute for Social Research, 426 Thompson St., Ann Arbor, Michigan 48106-1248. meganpat@umich.edu

Data collection and manuscript preparation were supported research grants from the National Institute on Drug Abuse (R01DA001411 and R01DA016575). The study sponsors had no role in the study design, collection, analysis or interpretation of the data, writing of the manuscript, or the decision to submit the paper for publication. The content is solely the responsibility of the authors and does not necessarily represent the official views of the study sponsor.

Recommended Citation

Patrick, M. E., Terry-McElrath, Y. M., Berglund, P., Pang, Y. C., Heeringa, S. G., & Si, Y. (2022). *An Updated Weighting Strategy for the Monitoring the Future Panel Study.* Monitoring the Future Occasional Paper No. 98. University of Michigan Institute for Social Research: Ann Arbor, MI. doi:10.7826/ISR-UM.06.585140.002.08.0001.2022

MonitoringtheFuture.org   @ysi_news   MTFinformation@umich.edu

UPDATED JULY 2023

An Updated Weighting Strategy for the Monitoring the Future Panel Study

Executive Summary

The Monitoring the Future (MTF) study monitors historical and developmental changes in substance use prevalence among key subgroups of the general U.S. adolescent and adult population. Begun in 1976, MTF's longitudinal panel study follows annual cohorts selected from nationally representative (i.e., national probability) samples of 12th grade students (modal age 18) into adulthood. As of the 2021 data collection, the panel study included longitudinal data from 45 cohorts, with the earliest cohorts now providing data at age 60. Addressing the impact of both the panel sample selection process and nonresponse is scientifically desirable to retain population representation and reduce attrition bias in the MTF panel study substance use estimates and trend analyses. The current study first devised and evaluated a cohort-specific pooled analysis weighting procedure for the MTF panel study that weighted back to the initial 12th grade samples, accounting for 1) the fraction of 12th grade students not eligible for panel selection, 2) the panel sample selection process including oversampling of those reporting drug use, and 3) panel attrition. Following this, the cohort-specific weights were updated to age-specific weights in order to provide increased flexibility for a wide range of model specifications and to avoid the need to re-calculate a first half-sample respondent's final weight at a particular wave after the full cohort had completed that wave.

Compared to prior weighting approaches (including the use of the traditional panel sampling weight currently in MTF datasets available through the National Addiction & HIV Data Archive Program, or NAHDAP), the cohort-specific pooled analysis weight approach that was developed resulted in 1) sociodemographic characteristic estimates at follow-up that more closely aligned with those from the original 12th grade samples and 2) slightly higher estimates of past 30-day cigarette and marijuana use and past 12-month use of drugs other than marijuana at the ages examined in the analyses (19/20 and 35) by key sociodemographic predictors (results for past 30-day alcohol use were somewhat mixed). Obtained standard errors were generally similar across weighting approaches. The cohort-specific pooled analysis weights appear to result in an overall improvement in the degree to which the sociodemographic distributions of the initial 12th grade samples are retained, as well as likely producing slightly improved substance use estimates due to accounting for historical variation in panel sample selection and attrition over time. The updated age-specific pooled analysis weights continued to provide the benefits associated with the cohort-specific weights, but also brought increased flexibility for modeling both cohort- and age-specific research questions, and allowed all respondents' weights at each specific age to remain fixed across time.

Users of MTF panel data should consider incorporating the age-specific pooled analysis weights developed by MTF analysts and available through the National Addiction & HIV Data Archive Program (NAHDAP). The age-specific pooled analysis weights should be relevant for analysis regardless of outcome focus (e.g., substance use, school and work success, health issues, participation in arts/sports programs, etc.). For analyses that combine data from multiple cohorts and/or ages, the pooled analysis weights should be normalized for the analytic sample.

Acknowledgments

The authors thank John S. Haeussler and Deborah D. Kloska for their contributions to this report. JS Haeussler provided a detailed review of the

information related to 12th grade sampling weights (Supplement Appendix B). DD Kloska provided a detailed review of the final report and assistance with recommendations and coding for weight selection and normalization (Supplement Appendix C).

MTF Panel data, MTF Panel analysis weights, MTF Restricted Panel DataUser's Guide, and accompanying documentation are available at: https://www.icpsr.umich.edu/web/NAHDAP/studies/37072

Introduction

The misuse of alcohol and other drugs in the United States (U.S.) is associated with staggering economic loss, societal harm, and premature loss of life (Recovery Centers of America, 2017); annual economic costs alone are estimated to be $249 billion for alcohol misuse and $193 billion for illicit drug use (Office of the Surgeon General, 2019). The nature of substance use is continually evolving: new substances are discovered and/or developed, new use methods become popular for existing substances, and substances that had been experiencing a decrease in use are rediscovered by new generations. Studies monitoring substance use patterns at the population level are critical for providing policymakers and prevention practitioners with information on national use trends and population subgroups at particularly high risk for harmful substance use.

One such study is Monitoring the Future (MTF), which provides the ability to monitor historical and developmental changes in substance use prevalence and trends, as well as related attitudes and beliefs among key general U.S. adult population subgroups, including college students and young adults (ages 19–30), and adults in midlife (ages 35–60). MTF is conducted at the University of Michigan's Institute for Social Research under a series of investigator-initiated, competing research grants from the National Institute on Drug Abuse (R01DA001411 and R01DA016575). The MTF study is based on several ongoing surveys, including the panel study which follows cohorts selected from annual nationally representative (i.e., national probability) samples of 12th grade students into adulthood. As of the 2021 data collection, the panel study included longitudinal data from 45 12th grade cohorts (1976–2020), following participants from modal ages 18 through 60 (with age 65 data collection starting in 2023).

The hallmark of the 12th grade MTF survey is its ability to provide nationally representative estimates of substance use prevalence and trends (for detailed methodology, see Bachman et al., 2015; Miech et al., 2022). Each

year, surveys are conducted in approximately 120–140[1] annually-selected public and private high schools to provide a national cross-sectional probability sample of 12th graders throughout the contiguous United States. A complex multistage probability sampling procedure is used for the 12th grade survey, with Stage 1 being the selection of particular geographic areas (with probability proportionate to the number of households), Stage 2 the selection of one or more high schools in each area (with probability proportionate to the student enrollment size for the 12th grade), and Stage 3 the selection of 12th graders within each high school. 12th grade sampling weights are assigned at the school level to compensate for differential probabilities of selection at each sampling stage and normalized to sum to the student sample size. It is the use of the resulting sampling weights that allows MTF 12th grade population estimates for a range of outcomes (including substance use) to be nationally representative. In turn, the obtained population estimates provide policy makers and prevention and intervention efforts with important information regarding substance use trends across cohorts and population subgroups; as these estimates change, both policy and prevention/intervention efforts can be modified to reduce risk. In analyses using panel data, it is scientifically desirable to retain the ability to obtain nationally representative estimates of substance use and associated trends that is possible when using the original 12th grade sample data, but doing so must address the impact of both the panel sample selection process and attrition. In other words, the number of individuals who are selected to participate in the panel study is much lower than the number initially participating in the 12th grade samples, so purposeful scientific revision to the original 12th grade sampling weights is necessary to retain national representation in panel analysis estimates. Without appropriate weights, obtained panel estimates cannot be said to be

[1] In 2020, the COVID-19 pandemic halted in-school data collection. See Johnston et al., 2022 and Miech et al., 2022 for further details on 2020 and 2021 data collection methods.

representative of the initial 12th grade national samples on which the panel is based and may lead to biased results due to selection bias.

Panel selection

The MTF panel study is based on selecting a cohort from each year's national 12th grade sample (who are surveyed at modal age 18). Beginning with the graduating class of 1976, from the 12,000–19,000 12th graders originally surveyed in a given year's sample, a cohort of 2,450 is chosen for follow-up in the panel study. The panel selection process impacts the representative nature of each 12th grade cohort in three main ways: 1) only 12th grade participants who provide contact information and non-missing data on sex[2] can be considered for panel selection, 2) those who report drug use are purposively oversampled, and 3) panel participants are selected with their probability inversely proportional to their cross-sectional 12th grade sampling weight. A traditional panel sampling weight addresses the oversampling of those reporting drug use, as well as panel selection probability inversely proportional to the 12th grade sampling weight. However, the traditional panel sampling weight does not address either 1) the proportion of original 12th graders who did not provide contact information and non-missing data on sex or 2) panel attrition.

Motivation for updating the panel weighting strategy

Currently, the traditional panel sampling weight is the only weight that has been provided on MTF panel datasets available through the National Addiction & HIV Data Archive Program (NAHDAP). The traditional panel sampling weight makes no adjustments for the increasing number of students not eligible for panel study participation due to not providing both contact information and/or data on sex (see Figure 1). Additionally, as is the case with all longitudinal data collection, study attrition over time is an

---

[2] Panel sampling includes equal selection by sex. On average, 5% of 12th grade respondents have missing data on sex. Valid data on sex was not required for the 12th grade cohorts of 2016–2018, as discussed below.

increasing concern. Survey response rates in general have been declining in recent years, and response is associated differentially with health risks including substance use (McCabe & West, 2016; Keyes et al., 2020; Si et al., 2022). Over time, unweighted MTF panel study response rates have ranged from a high of approximately 84% (at ages 19/20 for the 1980 cohort) to a low of approximately 33% (at ages 21/22 for the 2015 cohort) (see Supplement Table 2).

In recent years, MTF study investigator publications using panel data have incorporated wave-specific panel attrition weights calculated as the inverse of the response probability among participants selected for the panel study.[3] The response probability calculations have been based on past research that has identified key characteristics of 12th graders who provided contact information and were selected for the panel study. These characteristics are associated with panel response in substance use related analyses, including sex, race/ethnicity, parental education (used as a proxy for family of origin socio-economic status), number of parents in the home, average high school grades, truancy, college plans, substance use (alcohol, cigarettes, marijuana, and other drugs), the metropolitan status (i.e., large metropolitan, other metropolitan, or non-metropolitan) of the area surrounding the school, and U.S. Census region (Patrick et al., 2016; 2021b; Terry-McElrath et al., 2019; Terry-McElrath & O'Malley, 2015; Terry-McElrath & Patrick, 2016). However, traditional weights did not address the probabilities of providing either contact information or non-missing data on sex. No analyses had examined whether weighting procedures should model the probability of panel eligibility separately from panel attrition.

Supplemental analyses (reported in Supplement Appendix A) indicated the characteristics associated with the probability of panel response were also

---

[3] The wave-specific panel attrition weights used by MTF study investigators have not been available to NAHDAP users of the panel data; however, the User's Guide provides information and guidance on issues for researchers to consider with respect to accounting for panel attrition.

associated with panel eligibility (i.e., providing contact information and non-missing data on sex in 12th grade). The observed similarities are understandable in that the same measures are associated with the likelihood of participants providing any information. These similarities raised the question of whether the revised MTF panel analysis weighting procedure needed to separately model eligibility and response or whether both factors could be modeled jointly.

The current study

An effective approach for estimating substance use prevalence and trajectories in longitudinal research is using wave-specific inverse propensity score weighting (Si et al., 2022). In the case of MTF, an updated panel study analysis weighting approach that weights back to the national 12th grade samples should account for four key issues: 1) the probability of initial selection into the 12th grade samples, 2) the fraction of each 12th grade sample not eligible for panel selection because they did not provide contact information and/or data on sex, 3) the panel sample selection process including oversampling of those reporting drug use, and 4) panel attrition. All four of these propensities affect the ability of the panel samples to accurately reflect the original 12th grade national probability samples.

A primary question is if appropriate modeling requires a multistep process wherein panel eligibility, selection, and response propensities are modeled separately, or if a pooled modeling approach can be used that models these propensities simultaneously. The multistep approach would require one model predicting the probability of eligibility among the entire 12th grade sample and another model predicting wave-specific panel response probabilities among those selected for panel. The final multistep weight equation would involve four components (see Equation 1): 1) probability of selection into the 12th grade sample, 2) probability of eligibility for panel selection among all 12th grade participants, 3) probability of panel selection given eligibility, and 4) probability of wave-specific response. The advantage of using a pooled approach would lie in its comparative simplicity. Only one

analytic model using the entire 12th grade sample would be run predicting the probability of eligibility, selection, and wave-specific response, and it would be able to account for changes in panel sampling methodology over time by implicitly modeling selection simultaneously with eligibility and response. The final pooled weight calculation would involve only two components (see Equation 2): probability of selection into the 12th grade sample, and probability of eligibility, selection, and wave-specific response.

In order to provide new individual panel analysis weights for each wave/age of participation that allow for generalization to the original MTF 12th grade samples, the current study addressed three research aims:

Aim 1. Using data from two example panel cohorts (from 12th grade years 2000 and 2019), compare the effectiveness of cohort-specific panel analysis weights developed using 1) multistep modeling of panel eligibility, selection, and response propensities or 2) pooled modeling in which eligibility, selection, and response propensity are jointly estimated, with the goal of selecting the analysis weighting approach with optimum performance.

Aim 2. Using data from all panel cohorts from 1976 to 2020, compare the cohort-specific analysis weighting approach selected in Aim 1 with two alternative approaches: using the data unweighted and using the traditional panel sampling weight. Comparisons include examining the degree to which each approach reproduced baseline sociodemographic and substance use characteristics of the full original 12th grade samples—and examining prevalence estimates and standard errors at follow-up ages 19/20 and 35 for alcohol, cigarette, marijuana, and other drug use by sex, race/ethnicity, and parental education.

Aim 3. Update the cohort-specific pooled analyses weights to age-specific pooled analysis weights, thereby (a) increasing flexibility by addressing a wider range of analytic model specifications (e.g., accurately model both cohort- and age-specific models), and (b) preventing the need to re-calculate weights following the completion of a specified wave by a full cohort.

Methods

Detailed methods for the MTF panel study have been described elsewhere (Johnston et al., 2022; Miech et al., 2021a, 2022; Patrick et al., 2022a). Although the current analyses did not include the 2021 12th grade cohort (because panel data for that cohort is not yet available), this section provides a methodological overview for all cohorts up through 2021 to provide an understanding of recent changes relevant to sampling weight calculation.

12th grade data collection and panel eligibility

Data from 12th grade students are collected annually in the spring of each year. The 1976–2018 data collections involved in-school paper surveys from approximately 15,500 students per year (range 12,600–18,924); student response rates averaged 83% (range 77%–86%). In order to reduce respondent burden but still collect a wide range of measures, a total of six questionnaire forms are used (forms are randomly distributed within school). Data collection in 2019 involved in-school randomization to either paper or electronic tablets (Miech et al., 2021a); 13,713 students responded for an 80% response rate. In 2020, in-school data collection used electronic tablets only and was halted on March 15, 2020 because of the COVID-19 pandemic. The resulting 2020 sample size was approximately one-quarter the size of a typical data collection (3,770), with a 79% student response rate (Miech et al., 2021b, 2022). Data collection in 2021 utilized web-based surveys;[4] a total of 9,022 participants completed surveys remotely or in school for a response rate of 69%.

Per MTF study protocol, students were told that MTF would like to send them a summary of the study results, as well as possibly invite them to

---

[4] Web survey data collection for the 2021 12th grade sample occurred as 1) synchronous (the survey happened in "class time" usually with a teacher supervising either in person or remote), 2) asynchronous (where the survey link was given and assigned like homework), or 3) multiple (where surveys were both administered in class and as homework or done at the students' convenience).

complete another survey (i.e., entry to the panel study). Students were asked to provide their name and contact information: mailing address, phone number, and—depending on the year surveys were administered—cell phone number and email address.[5] Figure 1 provides trends in the total number of 12th grade students who participated by yearly sample and the percentage of students who 1) did not provide contact information, 2) did not provide data on sex, and 3) were not eligible for participation in the panel study due to either of the first two reasons. Increases over time in the percentage of students who were ineligible for panel selection have been driven primarily by increases in the number not providing contact information. These trends have been observed elsewhere and largely correspond to widely available email use in the mid-to late 1990s (Gibbs, 2016). For the 12th grade samples of 2016–2018, we explored the possibility of allowing missing data on sex for panel eligibility, but it was discovered that many of the participants with missing data on sex at 12th grade also provided little, if any, data at their first panel follow-up. Thus, from the 2019 12th grade sample onward, non-missing sex data were again required for panel study eligibility. The increase in the percentage of participants not providing contact information in 2019 coincided with the beginning of exploring electronic data collection methods for the 12th grade sample.

Decisions to select the 2000 and 2019 samples for Aim 1 in the current study were based on variation over time in panel eligibility and participation. The 2000 sample was selected because it occurred when the levels of panel ineligibility were still relatively low, and the 2019 sample was selected due to the dramatic increase in the missing contact information provision. Differences in panel response were also marked between the panel cohorts selected from these two samples (see Supplement Table 2). By choosing these two illustrative samples, analyses were able to explore the

[5] An "email address" was first requested in 2010; in 2020, this was changed to “non-school email." From 2016 onward, respondents were asked to provide two telephone numbers, one of which was labeled “your cell phone."

degree to which the two weighting approaches (multistep modeling vs. pooled modeling) were able to address historical differences in both eligibility and participation.

Panel selection and sampling weight

Historically, from each annual 12th grade sample, a cohort of about 2,450 participants has been selected for panel participation. The selection process begins by removing ineligible cases (those without contact information and/or data on sex) from consideration. Then, panel sample selection involves two main steps. First, all students are coded as "yes" or "no" for an indicator of past 30-day drug use.[6] Those in the drug use stratum are oversampled by a ratio of 3 to 1. Selection occurs within school by questionnaire form[7] and sex, and each 12th grade school is required to have a minimum of two panel selections (individuals). Important revisions to this process were needed for the 12th grade samples in 2020 and 2021. In 2020, 12th grade data collection was halted mid-study because of the COVID-19 pandemic (Miech et al., 2022). Due to the smaller number of cases available in the 12th grade sample, all eligible 2020 12th grade participants were selected with certainty for the panel cohort. From the 2021 12th grade sample, all eligible participants in the drug use stratum were selected with certainty (n=252), after which participants from the non-drug use stratum were subsampled in order to bring the cohort total to the desired 2,450 (2,198 in the non-drug use stratum plus 252 in the drug use stratum).

---

[6] Coded as a dichotomy indicating reporting use of marijuana on 20 or more occasions in the past 30 days; any past 30-day use of the illicit drugs cocaine, heroin, LSD or other hallucinogens; or any past 30-day use of the following without a doctor's prescription: amphetamines, sedatives or barbiturates, tranquilizers, or narcotics other than heroin.

[7] As noted above, there are currently six different questionnaire forms distributed to 12th grade students; forms are distributed randomly within school. For panel waves 1 (ages 19/20) to 6 (ages 29/30), respondents receive the same questionnaire form as they completed at 12th grade.

The traditional panel sampling weight accounts for three selection factors: 1) probability of 12th grade school selection, 2) proportion of 12th grade students in each sampled school that completed the survey, and 3) the conditional probability that an eligible 12th grade student was selected in the panel sample (i.e., oversampling based on 12th grade drug use stratum). The first two factors are reflected in the reciprocal of the 12th grade sampling weight. For the third factor, MTF has historically (through the 2019 12th grade cohort) assigned a constant factor of 1/3 for the drug use stratum and 1 for the non-drug stratum. For all 12th grade cohorts from 1976 to 2019, the traditional panel sampling weight had values of either 0.33333 or 1.0. This reflected the panel sample selection algorithm that subsampled cases with probability inversely proportional to their 12th grade sampling weight. Thus, in theory, multiplying the 12th grade sampling weight by the conditional panel sampling weight would result in the 12th grade sampling weight canceling out, leaving the conditional probability of panel selection based on 12th grade drug use status. In 2020 and 2021, the traditional panel sampling weight process was not applicable due to changes in the sample sizes of the 12th grade cohorts resulting from the COVID-19 pandemic. The usual "cancelling out" of the 12th grade sampling weight as part of panel sample selection was no longer a certainty for the 2020 cohort (when all eligible 12th grade participants were selected with certainty) or the 2021 cohort (when all eligible 12th graders who reported marijuana or other drug use were selected with certainty). Due to these changes, as well as recent revisions to the 12th grade sampling weight development process that have increased the range of 12th grade sampling weight values (see Supplement Appendix B), the range of values in the panel sampling weight for the 2020 cohort onward no longer follow the historical values of 0.33333 or 1.0.[8] The notable changes in the traditional panel sampling weight starting with the 2020 cohort have

[8] For the 2020 cohort, the panel sampling weight ranged from 0.37 to 11.0 with mean=1.0 and SD=1.05. For the 2021 cohort, the panel sampling weight ranged from 0.10 to 8.19, with mean=1.0 and SD=0.56.

been an additional factor driving the decision to update MTF panel study weighting procedures at this point.

Panel response

After selection into the panel study, participants are assigned to begin longitudinal follow-up either one year following 12th grade (age 19) or two years later (age 20). The half-sample assignment process accounts for drug use status, sex, survey form, and school. Each participant is surveyed every other year through age 29/30 (i.e., at ages 19/20, 21/22, 23/24, 25/26, 27/28, 29/30). Thus, there are six waves of data collection between ages 19 and 30 (known as follow-up surveys 1 to 6). By alternating the two random half samples from ages 19 through 30, the MTF panel study collects data from every cohort each year throughout young adulthood, even though any given individual participates every other year. Starting at age 35, participants are surveyed every 5 years with all cohort participants surveyed in the same year; currently these surveys continue through age 60 (see Supplement Table 1). As of the date of this original report (Summer 2022), only the first half-sample of the 2020 cohort had the opportunity to participate in the panel study at the follow-up 1 (age 19/20) in calendar year 2021.

Each calendar year, MTF panel data collection involves contacting approximately 28,500 participants across age and cohort. From 1977–2017, all panel data were collected using mailed questionnaires. In 2018 and 2019, a random half of those ages 19 to 30 were assigned to a web-push procedure, wherein contact and data collection procedures pushed participants to complete a web-based survey, with a mailed paper survey provided only for non-responders or if requested (Patrick et al., 2021a). In 2020, all surveys at ages 19 to 30 followed the web-push procedure; a random 50-50 split of paper versus web push was used to test the impact of the web-push procedures with those ages 35 to 60 (Patrick et al., 2022a). From 2021 onward, all panel data collection has used the web-push procedures. A University of Michigan institutional review board approved all aspects of the study.

Measures

12th grade administrative variables. The *12th grade sampling weight* was a continuous variable. Variables included due to their relevance to initial 12th grade school selection were *region* (U.S. Census definitions of Northeast, Midwest, South, and West), *metropolitan status* (large metropolitan area, other metropolitan area, or non-metropolitan area), and *school type* (public school, private religious school, or other private school). An additional school-level survey implementation variable, *tablet randomization*, was needed and relevant only for the 2019 sample and indicated randomization to electronic tablet condition (vs. paper survey condition). Finally, *form* indicated which of the six 12th grade questionnaire forms a participant had been randomly assigned to complete.

12th grade self-report measures. *Sex* was coded as female or male. *Race/ethnicity* was coded using mutually exclusive categories as Asian, Black, Hispanic, White, or Other/multiracial. *Parental education* indicated if at least one parent had a college degree (vs. not). *Number of parents* in the household was coded as two (vs. none/one). *Average high school grades* was a trichotomy of C+ or below, B- to B+, or A- or A. *Truancy* indicated the number of whole days of school missed because the student skipped or "cut" in the past four weeks (none, 1, 2, 3, 4 to 5, 6 to 10, 11 or more). *College plans* was coded as definitely plan to graduate from a 4-year college (vs. other). Past 30-day *cigarette use* was measured as not at all, <1/day, 1 to 5/day, about ½ pack/day, about 1 pack/day, about 1 ½ packs/day, 2+ packs/day. Past 30-day use of *alcohol* and *marijuana* were each measured using the same 7-point scale of 0 occasions, 1–2, 3–5, 6–9, 10–19, 20–39, 40+ occasions. *Drug use sampling indicator* was a binary measure used for oversampling into the panel, with drug use defined as using marijuana on 20 or more occasions in the past 30 days, or any non-medical use of other substances in the past 30 days: amphetamines, cocaine, heroin, other narcotics (opioids), hallucinogens, sedatives/barbiturates, or tranquilizers. *Providing contact information* indicated whether students provided both valid name and postal mailing addresses in 12th grade. Participants were considered *eligible for inclusion in*

the panel study if they provided contact information and data on sex, except for the 2016–2018 samples when missing data on sex was allowed.

Panel administrative measures. *Wave-specific response* indicated whether participants responded (yes or no) at each relevant wave (e.g., follow-up 1, 2, 3); participants were coded as missing for waves that were not yet relevant. Age-specific response indicated whether participants responded (yes or no) at each specific age (e.g., 19, 20, 21, etc.); participants were coded as missing for ages that were not yet relevant. For follow-up data collected at ages 19–30 in calendar years 2018–2019, and for data collected at ages 35–60 collected in calendar year 2020, *web-push randomization* indicated if a participant was randomized to web-push or standard mailed MTF data collection methods.

Panel self-report measures. Past *30-day cigarette, alcohol,* and *marijuana use* questions were the same as described above for 12th grade. Additionally, *use of drugs other than marijuana* was coded as a dichotomy to indicate any past 12-month non-medical use of amphetamines, cocaine, heroin, other narcotics (opioids), hallucinogens, sedatives/barbiturates, or tranquilizers.

Analyses

Analyses used SAS 9.4. Missing data on 12th grade covariates were addressed by using sequential imputation models to complete a single imputation by yearly sample (see Supplement Appendix A for further details). The imputed covariates were included in all regression models used for calculating probabilities utilized in weight construction.

Aim 1. For Aim 1 (comparing cohort-specific panel analysis weights obtained using multistep modeling versus pooled modeling), data from the 2000 and 2019 12th grade samples were used. Aim 1 analyses followed three steps: 1) calculate cohort-specific analysis weights for the follow-up 1 (age 19/20) wave using multistep modeling, 2) calculate analysis weights for follow-up 1 using pooled modeling, and 3) compare multistep and pooled weights.

Step 1: Calculation of cohort-specific multistep panel analysis weight ($PAW_{multistep}$). The multistep process accounted for initial 12^{th} grade selection, and then separately modeled panel eligibility, conditional selection, and conditional wave response, resulting in four weight components: 1) probability of selection for the 12^{th} grade sample (p_12^{th}), 2) probability of eligibility for panel selection among all 12^{th} grade participants (p_eligible), 3) probability of panel selection given eligibility (p_select), and 4) probability of wave-specific response at time t conditional on both eligibility and selection ($p\_response_t$) (see Table 1). The final calculation is shown in equation (EQ) 1:

EQ 1: cohort-specific $PAW_{multistep} = [(p\_12^{th}) \times (p\_eligible) \times (p\_select) \times (p\_response_t)]^{-1}$

Since eligibility was determined on a case-by-case basis,[9] all 12^{th} grade sample participants were used to model the probability of being eligible for panel selection based on selected covariates. Our argument for adding an explicit model of eligibility was that the mechanism for eligibility may differ substantially from that for response among eligible students (i.e., the characteristics associated with the probability of providing contact information in 12^{th} grade and therefore being eligible for selection may differ from the characteristics that predict response to the panel survey at any particular wave among selected individuals). We also included an explicit model of panel selection conditional on eligibility in order to account for previously described sample-specific changes in the constraints on the number of eligible cases (see Panel selection and sampling weight, above).

The MTF datasets provide the inverse of the probability of selection into the 12^{th} grade sample $(p\_12^{th})^{-1}$ as the 12^{th} grade sampling weight (see Supplement Appendix B). In order to calculate p_eligible by year among all the full 12^{th} grade samples for 2000 and 2019, a first logistic regression

[9] Eligibility defined from 1976–2015 and 2019 onwards as a respondent providing both valid contact information and non-missing data on sex; defined from 2016–2018 as a respondent providing valid contact information only.

model was used to regress the binary outcome of panel eligibility (eligible=1, otherwise=0) on the following: sex, race/ethnicity, parental education, number of parents, average high school grades, truancy, college plans, past 30-day cigarette use, past 30-day alcohol use, the drug use sampling indicator, school type, metropolitan status, and region (as well as a two-way interaction of metropolitan status and region). The 2019 model also included the tablet randomization covariate. To calculate *p*_select, by year among only those eligible for panel selection, a second logistic regression model was used to regress the 0/1 outcome of selected for panel on the following: sex, the drug use sampling indicator, the 12^{th} grade sampling weight, questionnaire form, and two-way interactions between the drug use sampling indicator and the three other covariates. To calculate $p\_response_t$ by cohort among those eligible and selected for the panel, a third logistic regression model was used to regress the 0/1 outcome of response at the specified wave (e.g., follow-up 1) on the same covariates used for modeling *p*_eligible: sex, race/ethnicity, parental education, number of parents, average high school grades, truancy, college plans, past 30-day cigarette use, past 30-day alcohol use, the drug use sampling indicator, school type, metropolitan status, and region (models for the 2019 cohort also included tablet randomization).

Step 2: Calculation of cohort-specific pooled panel analysis weight (PAW_{pooled}). The pooled process combined modeling of eligibility, selection, and wave-specific response, resulting in two weight components: 1) probability of selection for the 12^{th} grade sample (p_12^{th}) and 2) probability of eligibility, selection, and wave-specific response at time *t* ($p\_responseC_t$) (see Table 1). The final calculation is shown in equation (EQ) 2:

EQ 2: cohort-specific $PAW_{pooled} = [(p\_12^{th}) \times (p\_responseC_t)]^{-1}$

In this approach, we implicitly modeled the probability of an individual being sampled; we could not condition on inclusion when we jointly modeled eligibility, selection, and response as one process. In other words, the models included the full 12^{th} grade samples; the probability of selection

into the panel was not calculated only among those eligible for panel, and the probability of panel response was not calculated only among those selected for panel. The key variables related to the probability of selection into the panel were included in the models, along with measures associated with eligibility and response. To calculate p_responseC_t, by year among all the full 12th grade samples for 2000 and 2019, one logistic regression model was used to regress the 0/1 outcome of panel response at the specified wave (e.g., follow-up 1) on 1) the same covariates used for modeling p_eligible: sex, race/ethnicity, parental education, number of parents in the home, average high school grades, truancy, college plans, past 30-day cigarette use, past 30-day alcohol use, the drug use status, school type, metropolitan status, and region and 2) the 12th grade sampling weight, questionnaire sampling indicator, and two-way interactions between the drug use sampling indicator and sex, race/ethnicity, and metropolitan status. The model for the 2019 sample also included tablet randomization as a predictor.

Step 3: Analyses comparing cohort-specific $PAW_{multistep}$ and PAW_{pooled} included examining sums of weights, weight value coefficients of variation (defined as the ratio of the standard deviation to mean values), correlations between the two weights using PROC CORR, and using PROC SURVEYMEANS to examine 1) the degree to which the distributions and variances of sociodemographic characteristics of the original 12th grade samples were reproduced using weighted data from those in panel cohorts responding at follow-up 1 (age 19/20) and 2) differences in estimates of alcohol and marijuana use prevalence at follow-up 1.

Aim 2. For Aim 2 (demonstrating use of the cohort-specific PAW_{pooled} analysis weighting approach selected in Aim 1 compared with either using the data unweighted or using the traditional panel sampling weight), data from the 1976–2020 12th grade samples were used to calculate PAW_{pooled} by cohort and wave for all MTF panel participants as of 2021 data collection. Analyses included indicators for data collection changes as described above for relevant years, including tablet randomization (2019 cohort) and web-

push randomization (for cohorts responding at follow-ups 1 to 6 [ages 19/20 to 29/30] in calendar years 2018 and 2019 and for cohorts responding at ages 35–60 in calendar year 2020). For cohorts that had only their first half-sample completing a specific panel wave in 2021, models excluded cases selected for panel participation but scheduled to complete the specific wave the following year. For example, for data collected in calendar year 2021, only the first half-sample of the 2020 12th grade panel cohort was scheduled to be surveyed at age follow-up 1 (19/20). Thus, models predicting $p(\text{responseC}_{19/20})$ in 2021 excluded participants from the 2020 cohort who were not scheduled to complete the follow-up 1 survey until 2022. After calendar year 2022 data collection, all participants from the 2020 cohort will have had the opportunity to participate in the follow-up 1 survey. At that point, the entire 2020 sample (both those scheduled to participate at age 19 and at age 20) will be included in the same model predicting $p(\text{responseC}_{19/20})$.

We then compared the use of cohort-specific $\text{PAW}_{\text{pooled}}$ with both unweighted analyses and those using the traditional panel sampling weight by using PROC SURVEYMEANS to 1) reproduce the sociodemographic and substance use distributions of the combined original 12th grade samples and 2) obtain prevalence estimates and standard errors for alcohol, cigarette, marijuana, and other drug use among panel cohort participants at follow-up 1 (ages 19/20) and 35 by sex, race/ethnicity, and parental education.

Aim 3. For Aim 3 (updating the pooled weights to be age-specific vs. cohort-specific), the equation used to calculate cohort-specific $\text{PAW}_{\text{pooled}}$ was revised to focus on age-specific response. Specifically, the second weight component was revised to be the probability of eligibility, selection, and age-specific response (p_responseCA$_t$). The final calculation is shown in equation (EQ) 3:

EQ 3: age-specific $\text{PAW}_{\text{pooled}} = [(p\_12^{\text{th}}) \times (p\_\text{responseCA}_t)]^{-1}$

Data from the 1976-2020 12th grade samples were used to calculate age-specific $\text{PAW}_{\text{pooled}}$ by cohort and age for all MTF panel participants as of 2021

data collection. Analyses included the same indicators for data collection changes as described above for relevant years, including tablet randomization (2019 cohort) and web-push randomization for cohorts responding at follow-ups 1 to 6 at ages 19/20 to 29/30 in calendar years 2018 and 2019, and for cohorts responding at ages 35–60 in calendar year 2020.

The MTF panel study use of a half-sample approach for ages 19–30 was employed in order to allow one-half of each 12th grade cohort to represent the full cohort at a specified age, with two ages representing a single follow-up (e.g., ages 19 and 20 combined to make follow-up 1). As the half-sample selection process accounts for drug use status, sex, survey form, and school, each half-sample stands as a representative subsample of the original 12th grade sample, as does the full cohort (both half-samples together). This fact allows age-specific weights for each half-sample to be calculated that weight back to the full 12th grade sample.

There are two key differences in age-specific weights compared with cohort-specific weights. First, at each age from 19–30, models *do not* exclude cases selected for panel participation but scheduled to complete a specific wave the following year. For example, for data collected in calendar year 2021, only the first half-sample of the 2020 12th grade panel cohort was scheduled to be surveyed at age 19, but models predicting $p(\text{responseCA}_{19})$ in 2021 included all 2020 cohort participants. Second, the resulting weight becomes the permanent pooled analysis weight at age t for these respondents, and does not need re-calculation after a cohort's second half-sample completes the specified follow-up. Since surveys at ages 35–60 involve surveying the entire cohort in the same year, age-specific and cohort-specific weights at these ages are identical.

Results

Aim 1: Cohort-specific multistep versus pooled analysis weights at follow-up 1 (age 19/20) for 2000 and 2019 panel cohorts

A total of 13,286 participants were included in the 2000 12th grade sample; of the 2,456 selected for the 2000 panel cohort, 1,429 participants responded at follow-up 1 (age 19/20) in calendar years 2001 and 2002. The total 2019 12th grade sample included 13,713 participants; of the 2,447 selected for the 2019 panel cohort, 725 responded at follow-up 1 in calendar years 2020 and 2021. Table 2 presents, by 12th grade sample and the drug use sampling indicator, mean predicted probabilities of panel eligibility, selection, and response at follow-up 1 following the cohort-specific $PAW_{multistep}$ method. Results show that, in general, participants in the drug use stratum were less likely to be eligible for the panel (i.e., provide both contact information and data on sex), more likely to be selected for panel (due to purposeful oversampling based on drug use), and less likely to respond at follow-up.

Hosmer and Lemeshow Chi-square Goodness of Fit (HL GoF) tests for models predicting probability of response showed reasonable fit; *p*-values for the 2000 cohort were 0.418 for $p\_response_{19/20}$ (used in cohort-specific $PAW_{multistep}$) and 0.517 for $p\_responseC_{19/20}$ (used in cohort-specific PAW_{pooled}). Respective HL GoF *p*-values for the 2019 cohort were 0.144 and 0.523. Table 3 provides descriptive statistics for cohort-specific $PAW_{multistep}$ and PAW_{pooled} at follow-up 1 (age 19/20) for the 2000 and 2019 cohorts. For both cohorts, the sum of weights for cohort-specific PAW_{pooled}—a check on the accuracy of the model and weight computation process—was closer to the original 12th grade sample size than the sum of weights for cohort-specific $PAW_{multistep}$. For the 2000 cohort, the coefficient of variation for cohort-specific $PAW_{multistep}$—a measure of inflated variability in the computed weights—was slightly lower than cohort-specific PAW_{pooled} (0.52 vs. 0.55), but for the 2019 cohort, the coefficient of variation for cohort-specific $PAW_{multistep}$ was noticeably higher than cohort-specific PAW_{pooled} (0.86 vs. 0.71). The

correlation between cohort-specific $PAW_{multistep}$ and PAW_{pooled} for the 2000 cohort was 0.93 ($p<.001$) and for the 2019 cohort was 0.91 ($p<.001$).

Table 4 presents weighted descriptive statistics of 12th grade covariates for 1) the entire 12th grade samples (weighted using the 12th grade sampling weight) and 2) those in panel cohorts responding at follow-up 1 (age 19/20) weighted first with cohort-specific $PAW_{multistep}$ and then with cohort-specific PAW_{pooled}. Standard errors at follow-up 1 were quite similar when comparing cohort-specific $PAW_{multistep}$ and PAW_{pooled} for the 2000 cohort but were generally smaller when using cohort-specific PAW_{pooled} for the 2019 cohort. Differences in 12th grade descriptive statistic estimates between the entire 12th grade samples and those responding in panel cohorts at follow-up 1 were, overall, slightly smaller for the 2000 cohort and slightly larger for the 2019 cohort when using cohort-specific PAW_{pooled} (vs. $PAW_{multistep}$). However, differences were relatively small in magnitude. For example, the largest difference for a dichotomous outcome was observed for parental education for the 2000 sample (proportion = 0.516 for the entire sample vs. 0.527 and 0.524 for the panel cohort at follow-up 1 using cohort-specific $PAW_{multistep}$ and PAW_{pooled}, respectively) and for having 2 or more parents in the home for the 2019 sample (proportion = 0.641 for the entire sample vs. 0.632 and 0.653 for the panel cohort at follow-up 1 using cohort-specific $PAW_{multistep}$ and PAW_{pooled}, respectively).

Table 5 provides comparative estimates of past 30-day alcohol and marijuana use prevalence at follow-up 1 (age 19/20) by panel cohort. For both alcohol and marijuana use, estimates showed overlapping 95% confidence based on use of cohort-specific $PAW_{multistep}$ and PAW_{pooled}.

After evaluation of the above findings, cohort-specific PAW_{pooled} was identified as the preferred weighting approach for calculating the final panel analysis weights in Aim 2 due to 1) sums of weights that were closer to the original 12th grade sample sizes, 2) very similar results for both coefficient of variation and weighted 12th grade descriptive statistic estimates and standard errors for the 2000 panel cohort, and 3) a noticeably lower

coefficient of variation and somewhat smaller differences in descriptive statistic standard errors for 12th grade covariate estimates for the 2019 panel cohort.

Aim 2: Demonstration of the cohort-specific PAW_{pooled} analysis weight using 1976–2020 cohorts

Table 6 provides descriptive statistics for the cohort-specific PAW_{pooled} values for panel cohort participants from the 1976–2020 samples. As expected, overall, the longer the term of follow-up, the greater the cumulative effects of attrition and the larger the mean individual weight needed to compensate for cumulative attrition. For example, mean cohort-specific PAW_{pooled} was 10.5 at follow-up 1 (ages 19/20), and 14.4 at age 35. However, as can be seen from the table, some degree of decrease in mean individual weights was observed among the later age waves. For example, mean cohort-specific PAW_{pooled} was 15.3 at age 50, 15.1 at age 55, and 14.7 at age 60. This decrease is attributed to markedly lower attrition rates observed for the early 12th grade cohorts across panel follow-up.

Tables 7a and 7b provide percentages and standard errors for key 12th grade sociodemographic variables, providing comparative data for 1) combined original 12th grade samples and 2) participants in the panel cohorts. The two follow-ups of interest were follow-up 1 (ages 19/20) and age 35. As of 2021 data collection, full cohorts from the 1976–2019 12th grade samples, as well as the first half-sample of the 2020 cohort, had been able to participate at follow-up 1. Full cohorts from the 1976–2004 12th grade samples had been able to participate at the age 35 wave. Thus, Table 7a provides estimates for the noted variables for all 1976–2020 12th grade samples combined, and then for those participants from the 1976–2020 panel cohorts who responded at follow-up 1. Table 7b provides estimates for the noted variables for all 1976–2004 12th grade samples combined, and then for those participants from the 1976–2004 panel cohorts who responded at age 35. Four weighting strategies are used in these tables: 12th grade sampling weight with full 12th grade samples, and then three separate strategies for

the panel cohorts: 1) unweighted, 2) traditional panel sampling weight, and 3) cohort-specific PAW_{pooled}. Evaluation of the resulting estimates suggests that cohort-specific PAW_{pooled} performed well. Compared to results from the full 12th grade samples weighted with the 12th grade sampling weight, estimates obtained for the panel cohorts using PAW_{pooled} at both follow-up 1 and age 35 appeared to very closely align with those from the full 12th grade samples. Among the panel cohorts, unweighted estimates and those using the traditional panel sampling weight appeared to slightly underrepresent the following subgroups: males; those who identified as Black, Hispanic, or Other race; those who did not have at least one parent with a college degree; and those in the South and West in 12th grade.

Tables 8a and 8b present prevalence estimates of past 30-day alcohol, cigarette, and marijuana use, along with past year other drug use, by sex, race/ethnicity, and parent education among panel cohort participants responding at follow-up 1 (ages 19/20; cohorts 1976–2020) and age 35 (cohorts 1976–2004). These tables compare unweighted and weighted estimates using the traditional panel sampling weight and cohort-specific PAW_{pooled} for the noted follow-up waves. In general, unweighted substance use estimates across cohort population subgroups were somewhat higher than those obtained using either the traditional panel sampling weight or cohort-specific PAW_{pooled}. Compared with the traditional panel sampling weight, use of cohort-specific PAW_{pooled} consistently produced 1) somewhat higher estimates (with similar standard errors) for cigarette, marijuana, and other drug use; 2) somewhat lower estimates for alcohol use at follow-up 1; and 3) mixed results for alcohol use at age 35. The slightly higher estimates for cigarette, marijuana, and other drug use may be explained by the slightly greater upweighting of participants in the 12th grade drug user stratum relative to those in the non-user stratum. As noted in Table 2, those in the 12th grade drug user stratum had smaller response probabilities than those in the non-user stratum (i.e., were more likely to drop out of the study).

Aim 3: Updating the pooled analysis weights to be age-specific versus cohort-specific

Table 9 provides descriptive statistics for the age-specific PAW_{pooled} values for panel cohort participants from the 1976–2020 samples. There is a noticeable difference in mean values between cohort-specific PAW_{pooled} for follow-ups 1 to 6 and age-specific PAW_{pooled} for ages 19 to 30. For example, the cohort-specific mean PAW_{pooled} for follow-up 1 (age 19/20) was 10.5, while the age-specific mean PAW_{pooled} was 20.3 for age 19 and 21.5 for age 20. These differences result from how many panel respondents are included when weighting back to the original 12th grade samples. For cohort-specific weights, the entire cohort (approximately 2,450 individuals) is used to weight back to the initial 12th grade sample. In contrast, for age-specific weights, the first and second half-samples independently weight back to the initial 12th grade sample, thereby resulting in mean values roughly doubling for age-specific weights. Since the full cohort is sampled in the same year for ages 35 and up, mean age-specific and cohort-specific PAW_{pooled} values are equal for these ages. As with cohort-specific PAW_{pooled}, for age-specific $PAW_{pooled,}$ the longer the term of follow-up, the greater the cumulative effects of attrition and the larger the mean individual weight needed to compensate for cumulative attrition. For example, mean age-specific PAW_{pooled} was 20.3 at age 19, and 26.8 at age 30.

Following the development of the age-specific PAW_{pooled} weights, extensive analyses were conducted to examine the degree to which results varied when using age-specific versus cohort-specific weights. Specifically, all point estimates and trend comparisons reported in the MTF annual report on young adult and adult data (Patrick et al., 2022b) were re-run using both cohort-specific and age-specific PAW_{pooled} weights; the results were compared with the published estimates that used the traditional panel sampling weights. Results indicated that both cohort-specific and age-specific weights produced similar results. The percentage of all point estimate comparisons that exhibited less than a |5%| mean change between weighting approaches was 61.7% for cohort-specific weights and 63.1% for

age-specific weights. No trend comparisons indicated a sign flip, and the percentage that was substantively unchanged in regards to trend significance was 88.9% for cohort-specific weights and 87.1% for age-specific weights. Due to advantages related to use flexibility and permanence (discussed below), age-specific weights were identified as the preferred weighting approach and used in the final report comparing use of pooled weighting and traditional panel sampling weights (Terry-McElrath & Patrick, 2023).

Age-specific PAW_{pooled} weights were developed to address two main issues (compared with cohort-specific PAW_{pooled} weights):

- First, it is an advantage to have a single weight that is flexible enough to address both cohort-specific and age-specific research questions. Until age 35, the full cohort from any specified 12th grade school is never surveyed in the same year (as they are split by half-sample). This fact becomes relevant when considering the correct analytic dataset used for a specific research question. If a research question requires the analytic dataset to include the full cohort across follow-ups (e.g., prevalence of a specific substance for the full 2005 and 2006 12th grade cohorts across young adulthood [follow-ups 1 to 6]), then both cohort-specific and age-specific weights work well. For these situations, models using cohort-specific weights are conceptually combining the two ages included in each follow-up as a single data collection, with the need to also control for age. Models using age-specific weights are conceptually viewing each age as independent.

  However, other types of research questions specifically require the use of half-samples, such as (but not limited to):

  - analyses that involve only one year of data collection (e.g., the prevalence of a specific drug in 2018 across ages);

  - analyses that focus on a particular age across historical time (e.g., prevalence of a specific drug at age 19 from 1976 to 2020);
  - analyses that focus on multiple years of data collection but do not require that a full cohort has completed the relevant follow-up, such as prevalence for a specific drug in 2019 and 2020 across ages 19 to 26. This results in having only the first half-sample of the 2019 cohort (age 19, surveyed in 2020), the full 2018-2012 cohorts, and only the second half-sample of the 2011 cohort (age 26, surveyed in 2019).
- Second, it is an advantage to have all weights be permanent and not require re-calculation. Calculation of the cohort-specific weights required that weight values for the first half-sample of a cohort at a specific follow-up be re-calculated after the second half-sample completed the relevant follow-up. Changing weight values for respondents over time can lead to issues with replicability and increases burden for analysts.

The updated age-specific PAW_{pooled} weights (1) weight back to the original 12^{th} grade samples, (2) can be used in all types of analytic datasets, whether a full cohort or half-sample is being used for a particular 12^{th} grade sample, and (3) are permanent for each respondent.

Conclusion

The current study set out to (1) compare the effectiveness of multistep versus pooled MTF panel analysis weights (accounting for panel eligibility, selection, and response propensities) and select the weighting approach with better performance; (2) compare how the selected weighting approach performed in actual analyses with two alternative approaches using MTF panel data (using the data unweighted or using the traditional panel sampling weight); and (3) update the selected pooled weighting approach to be age-specific versus cohort-specific.

Predictors of eligibility for panel selection were the same as those for panel response (i.e., in the model of eligibility and in the separate model of panel selection conditional on eligibility). Analyses indicated that a pooled modeling approach that used these predictors with all possible 12^{th} grade sample participants to predict the probability of response at cohort-specific follow-ups was able to successfully weight back to the 12^{th} grade national samples while accounting for: 1) the fraction not invited to panel because they did not provide contact information or data on sex, 2) the panel sample selection process including oversampling of those reporting drug use, and 3) panel attrition. The pooled modeling approach (PAW_{pooled}) appeared to successfully address changes in 12^{th} grade sampling weight strategies and both 12^{th} grade and panel survey administration methodology (e.g., moving from paper to tablet at 12^{th} grade, and from paper to web-push procedures at follow-up), as well as the impact of unforeseen events such as the COVID-19 pandemic that significantly affected data collection efforts in 2020 and 2021. The update to age-specific pooled analysis weights increased flexibility by addressing a wider range of analytic model specifications (e.g., accurately model both cohort- and age-specific models), and (b) prevented the need to re-calculate weights following the completion of a specified wave by a full cohort.

We anticipate the age-specific pooled modeling approach will also be able to address future changes as they occur. The new age-specific PAW_{pooled} weights should be generally useful for all types of analyses (with varied outcomes such as substance use, school and work success, health issues, and participation in arts/sports programs). The weights are developed by the MTF research team and should be applied universally to all longitudinal analyses.

Recommendations and considerations for use

- Analyses using the MTF panel study data should include PAW_{pooled} weights (available at https://www.icpsr.umich.edu/web/NAHDAP/studies/37072). Given that

the PAW_{pooled} weights do not present extreme values, trimming is not needed. The PAW_{pooled} process produces a single panel analysis weight per participant for each follow-up age, with up to 18 weights possible at present (up through age 60). Appropriate decisions on weight selection will depend on the structure of the data ("wide" vs. "skinny/long" datasets), the type of modeling being done (e.g., multi-level modeling that can include two-level weights), and the degree to which participants included in analyses have had the opportunity to reach the same follow-up age or wave. Guidance on weight selection can be found in Supplement Appendix C.

- If a particular research question involves combining data from multiple cohorts or ages, it is advisable to normalize the PAW_{pooled} weights for the analytic dataset (i.e., divide each participant's PAW_{pooled} value by the mean for the analytic dataset) so that each cohort half-sample contributes to the analyses in proportion to its unweighted *n*. Guidance on weight normalization can be found in Supplement Appendix C. Normalization is recommended because the PAW_{pooled} weights have been developed to weight back to the 12th grade national probability samples. As described previously, each panel cohort has an approximate size of 2,450 (with each half-sample having an approximate size of 1,225), while each yearly 12th grade sample has an approximate size of 12,000–19,000. Using the PAW_{pooled} weights in analyses results in a weighted *n* that reflects the number of 12th grade students in the original samples that are represented by the panel cohorts, which is much larger than the cohort sample sizes.

- For analyses using form-specific data (i.e., based on data only included on some of the randomly-distributed survey forms), users should proceed with use of the PAW_{pooled} weights in the same manner as for the full sample, including normalizing procedures. As not all measures are included on all forms for those ages 19–30, some analyses will be form-specific (i.e., limited to respondents who responded to the form(s) of interest).

References

Bachman, J. G., Johnston, L. D., O'Malley, P. M., Schulenberg, J. E., & Miech, R. A. (2015). *The Monitoring the Future project after four decades: Design and procedures.* (Monitoring the Future Occasional Paper No. 82). Ann Arbor, MI: Institute for Social Research, University of Michigan. http://www.monitoringthefuture.org/pubs/occpapers/mtf-occ82.pdf

Gibbs, S. (2016, March 7). How did email grow from messages between academics to a global epidemic? *The Guardian*. https://www.theguardian.com/technology/2016/mar/07/email-ray-tomlinson-history

Johnston, L. D., Miech, R. A., O'Malley, P. M., Bachman, J. G., Schulenberg, J. E., & Patrick, M. E. (2020). *Demographic subgroup trends among adolescents in the use of various licit and illicit drugs, 1975-2019*. (Monitoring the Future Occasional Paper No. 94). Ann Arbor, MI: Institute for Social Research, University of Michigan. http://monitoringthefuture.org/pubs/occpapers/mtf-occ94.pdf

Johnston, L. D., Miech, R. A., O'Malley, P. M., Bachman, J. G., Schulenberg, J. E., & Patrick, M. E. (2022). *Monitoring the Future national survey results on drug use, 1975-2021: 2021 overview, key findings on adolescent drug use.* Ann Arbor: Institute for Social Research, The University of Michigan. http://monitoringthefuture.org/pubs/monographs/mtf-overview2021.pdf

Keyes, K. M., Jager, J., Platt, J., Rutherford, C., Patrick, M. E., Kloska, D. D., & Schulenberg, J. E. (2020). When does attrition lead to bias? Bias analysis for loss to follow-up in 30 longitudinal cohorts. *International Journal of Methods in Psychiatric Research, 29*(4), Article e1842. https://doi.org/10.1002/mpr.1842

McCabe, S. E., & West, B. T. (2016). Selective nonresponse bias in population-based survey estimates of drug use behaviors in the

United States. *Social Psychiatry & Psychiatric Epidemiology, 51*(1), 141-153. https://doi.org/10.1007/s00127-015-1122-2

Miech, R. A., Couper, M. P., Heeringa, S. G., & Patrick, M. E. (2021a). The impact of survey mode on US national estimates of adolescent drug prevalence: results from a randomized controlled study. *Addiction 116*(5), 1144-1151. https://doi.org/10.1111/add.15249

Miech, R. A., Johnston, L. D., O'Malley, P. M., Bachman, J. G., Schulenberg, J. E., & Patrick, M. E. (2022). Monitoring the Future national survey results on drug use, 1975-2021: Volume I, Secondary school students. Ann Arbor: Institute for Social Research, The University of Michigan. https://monitoringthefuture.org/wp-content/uploads/2022/08/mtf-vol1_2021.pdf

Miech, R. A., Patrick, M. E., Keyes, K. M., O'Malley, P. M., & Johnston, L. D. (2021b). Adolescent drug use before and during U.S. national COVID-19 social distancing policies. *Drug and Alcohol Dependence, 226*(September), 108822. https://doi.org/10.1016/j.drugalcdep.2021.108822

Office of the Surgeon General. (2019). Addiction and Substance Misuse Reports and Publications. https://www.hhs.gov/surgeongeneral/reports-and-publications/addiction-and-substance-misuse/index.html

Patrick, M. E., Couper, M. P., Parks, M. J., Laetz, V., & Schulenberg, J. E. (2021a). Comparison of a web-push survey research protocol with a mailed paper and pencil protocol in the Monitoring the Future panel survey. *Addiction, 116*(1), 191-199. https://doi.org/10.1111/add.15158

Patrick, M. E., Kloska, D. D., Mehus, C., Terry-McElrath, Y. M., O'Malley, P. M., & Schulenberg, J. E. (2021b). Key subgroup differences in age-related change from 18 to 55 in alcohol and marijuana use: U.S. national data. *Journal of Studies on Alcohol and Drugs, 82*(1), 93-102. https://doi.org/10.15288/jsad.2021.82.93

Patrick, M. E., Pang, Y. C., Terry-McElrath, Y. M., Laetz, V., & Couper, M. P. (2022a). Comparison of a web-push vs. mailed survey protocol in the Monitoring the Future panel study among adults ages 35 to 60. *Drug and Alcohol Dependence Reports, 4*(September), 100089. https://doi.org/10.1016/j.dadr.2022.100089

Patrick, M. E., Schulenberg, J. E., Miech, R. A., Johnston, L. D., O'Malley, P. M., & Bachman, J. G. (2022b). Monitoring the Future Panel Study annual report: National data on substance use among adults ages 19 to 60, 1976-2021. Monitoring the Future Monograph Series. Ann Arbor: Institute for Social Research, The University of Michigan.

Patrick, M. E., Terry-McElrath, Y. M., Kloska, D. D., & Schulenberg, J. E. (2016). High-Intensity drinking among young adults in the United States: Prevalence, frequency, and developmental change. *Alcoholism: Clinical and Experimental Research 40*(9), 1905-1912. https://doi.org/10.1111/acer.13164

Recovery Centers of America. (2017). Economic Cost of Substance Abuse in the United States, 2016. https://recoverycentersofamerica.com/PDF/Economic-Cost-of-Substance-Abuse.pdf

Schulenberg, J. E., Patrick, M. E., Johnston, L. D., O'Malley, P. M., Bachman, J. G., & Miech, R. A. (2021). *Monitoring the Future national survey results on drug use, 1975-2020: Volume II, college students and adults ages 19-60*. Ann Arbor: Institute for Social Research, The University of Michigan. http://www.monitoringthefuture.org/pubs/monographs/mtf-vol2_2020.pdf

Si, Y., West, B. T., Veliz, P., Patrick, M. E., Schulenberg, J. E., Kloska, D. D., Terry-McElrath, Y. M., & McCabe, S. E. (2022). An empirical evaluation of alternative approaches to adjusting for attrition when analyzing longitudinal survey data on adolescents' substance use trajectories. *International Journal of Methods in Psychiatric Research, 31*(3), e1916. http://doi.org/10.1002/mpr.1916

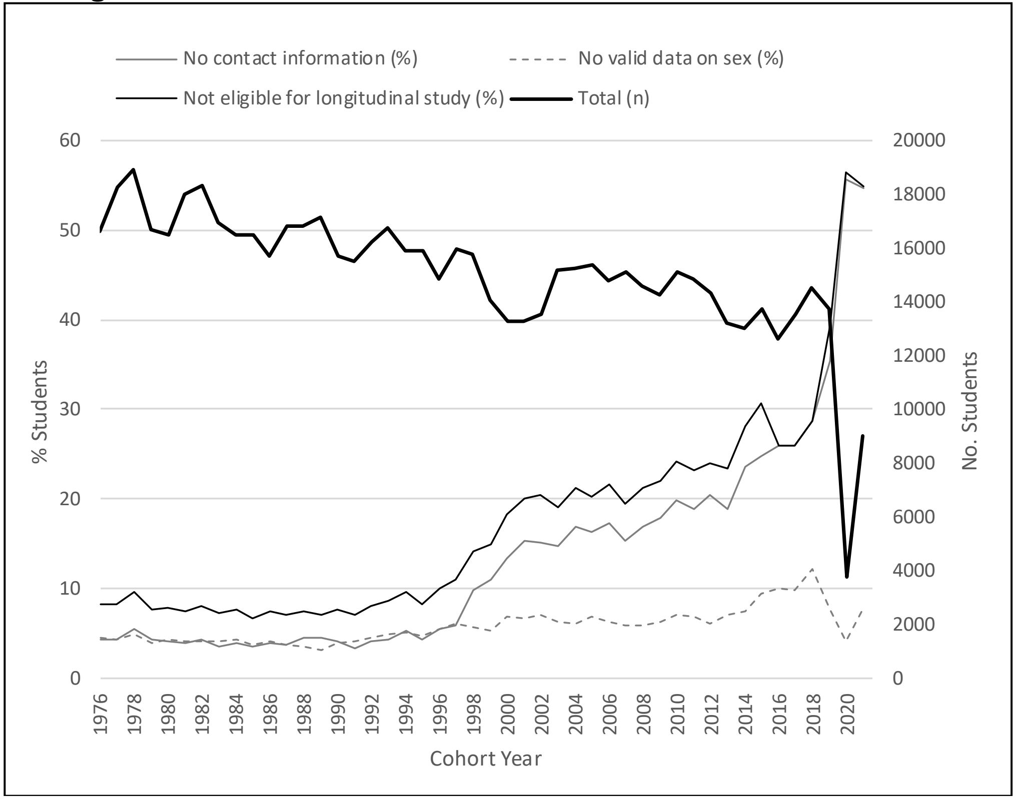
Terry-McElrath, Y. M., & Patrick, M. E. (2016). Intoxication and binge and high-intensity drinking among U.S. young adults in their mid-twenties. *Substance Abuse, 37*(4), 597-605. https://doi.org/10.1080/08897077.2016.1178681

Terry-McElrath, Y. M. & Patrick, M. E. (2023). Comparison of estimates before and after the updated weighting strategy change for the Monitoring the Future Panel study annual report. Monitoring the Future Occasional Paper No. 100. Ann Arbor, MI: Institute for Social Research, University of Michigan.

Terry-McElrath, Y. M., Patrick, M. E., Johnston, J. D., & Schulenberg, J. E. (2019). Young adult longitudinal patterns of marijuana use among U.S. National samples of 12th grade frequent marijuana users: a repeated-measures latent class analysis. *Addiction, 114*(6), 1035-1048. https://doi.org/10.1111/add.14548

Terry-McElrath, Y. M, & O'Malley, P. M. (2015). Trends and timing of cigarette smoking uptake among US young adults: survival analysis using annual national cohorts from 1976 to 2005. Addiction 110(7):1171-1181. https://doi.org/10.1111/add.12926

Figure 1. Trends in 12th grade student provision of contact information and non-missing data on sex, 1976–2021



Notes: From 2016–2018, missing data on sex was allowed for longitudinal study participation. From 1976–2020, data collection occurred in school. In 2020, data collection was halted mid-study due to the COVID-19 pandemic; data collection in 2021 occurred virtually using web surveys.

Table 1. Cohort-specific multistep and pooled panel analysis weight development

Cohort-Specific Multistep Panel Analysis Weight		
1.	p(12th grade sample participant)	(p_12th)
2.	p(eligible for panel | 12th grade participant)	(p_eligible)
3.	p(selected for panel | eligible 12th grade participant)	(p_select)
4.	p(respond at panel wave t | eligible, selected 12th grade participant)	(p_response$_t$)
	Multistep panel analysis weight (PAW$_{multistep(t)}$) = $[(p\_12^{th}) \times (p\_eligible) \times (p\_select) \times (p\_response_t)]^{-1}$	
Cohort-Specific Pooled Panel Analysis Weight		
1.	p(12th grade sample participant)	(p_12th)
2.	p(combined eligible, selected, and responded at panel wave t | 12th grade participant)	(p_responseC$_t$)
	Pooled panel analysis weight (PAW$_{pooled(t)}$) = $[(p\_12^{th}) \times (p\_responseC_t)]^{-1}$	

Table 2. Predicted probabilities of panel eligibility, selection and response at follow-up 1 (age 19/20) by drug use sampling indicator: 12th grade samples of 2000 and 2019

12th grade sample	Drug use sampling indicator	n	Mean	(SD)
			p(eligible)[a]	
2000	Non-user stratum	11,694	0.827	(0.087)
	User stratum	1,592	0.749	(0.115)
2019	Non-user stratum	12,463	0.616	(0.126)
	User stratum	1,250	0.517	(0.149)
			p(select)[b]	
2000	Non-user stratum	9,666	0.184	(0.082)
	User stratum	1,192	0.571	(0.263)
2019	Non-user stratum	7,676	0.256	(0.054)
	User stratum	646	0.754	(0.221)
			$p(\text{response}_{[19\text{-}20]})$[c]	
2000	Non-user stratum	1,775	0.598	(0.145)
	User stratum	681	0.540	(0.145)
2019	Non-user stratum	1,963	0.321	(0.127)
	User stratum	487	0.195	(0.098)

[a] *p*(eligible) = probability of eligibility for panel selection out of all participating students in noted 12th grade sample.
[b] *p*(select) = conditional probability of selection among eligible students in noted 12th grade sample.
[c] $p(\text{response}_{[19\text{-}20]})$ = conditional probability of response at follow-up 1 (age 19/20) among eligible, selected students in noted panel cohort.

Table 3. Descriptive summaries for cohort-specific multistep ($PAW_{multistep}$) and cohort-specific pooled (PAW_{pooled}) panel analysis weights at follow-up 1 (age 19/20): 12th grade samples from 2000 and 2019

12th grade sample	N[a]		Cohort-specific $PAW_{multistep(19/20)}$	Cohort-specific $PAW_{pooled(19/20)}$
2000	13,286	N participating at follow-up 1	1,429	1,429
		Sum of weights	13,817.6	13,769.2
		Mean (SD)	9.7 (5.1)	9.6 (5.3)
		Coefficient of variation	0.52	0.55
		Range	1.6 - 39.9	1.1 - 43.8
2019	13,713	N participating at follow-up 1	725	725
		Sum of weights	14,366.3	13,653.9
		Mean (SD)	19.8 (17.0)	18.8 (13.4)
		Coefficient of variation	0.86	0.71
		Range	1.9 - 212.4	1.9 - 103.0

[a] Total N of 12th graders in sample.

Table 4. Comparison of cohort-specific multistep and pooled panel analysis weight approaches at follow-up 1 (age 19/20): 12th grade cohorts of 2000 and 2019

	2000 12th grade sample			2019 12th grade sample		
	Full sample	Follow-up 1 panel cohort participants		Full sample	Follow-up 1 panel cohort participants	
	12th grade sampling weight Prop/Mean (SE)	Cohort-specific $PAW_{multistep}$ Prop/Mean (SE)	Cohort-specific PAW_{pooled} Prop/Mean (SE)	12th grade sampling weight Prop/Mean (SE)	Cohort-specific $PAW_{multistep}$ Prop/Mean (SE)	Cohort-specific PAW_{pooled} Prop/Mean (SE)
Sex						
Female	0.519 (0.005)	0.512 (0.015)	0.511 (0.015)	0.516 (0.005)	0.510 (0.025)	0.523 (0.023)
Male	0.481 (0.005)	0.488 (0.015)	0.489 (0.015)	0.484 (0.005)	0.490 (0.025)	0.477 (0.023)
Race/ethnicity						
Asian	0.039 (0.002)	0.039 (0.006)	0.039 (0.006)	0.037 (0.002)	0.044 (0.011)	0.053 (0.013)
Black	0.139 (0.004)	0.135 (0.011)	0.143 (0.012)	0.110 (0.003)	0.110 (0.017)	0.102 (0.014)
Hispanic	0.094 (0.003)	0.092 (0.009)	0.088 (0.009)	0.221 (0.004)	0.238 (0.024)	0.218 (0.020)
White	0.675 (0.005)	0.683 (0.015)	0.676 (0.015)	0.527 (0.005)	0.514 (0.025)	0.536 (0.023)
Other	0.053 (0.002)	0.051 (0.008)	0.055 (0.009)	0.105 (0.003)	0.094 (0.013)	0.092 (0.012)
Number of parents in the household						
Less than 2	0.300 (0.005)	0.288 (0.014)	0.300 (0.015)	0.359 (0.005)	0.368 (0.025)	0.347 (0.023)
Two	0.700 (0.005)	0.712 (0.014)	0.700 (0.015)	0.641 (0.005)	0.632 (0.025)	0.653 (0.023)
Parental education (at least 1 parent had college degree						
No	0.484 (0.005)	0.473 (0.015)	0.476 (0.015)	0.466 (0.005)	0.449 (0.025)	0.421 (0.022)
Yes	0.516 (0.005)	0.527 (0.015)	0.524 (0.015)	0.534 (0.005)	0.551 (0.025)	0.579 (0.022)
Average high school grades						
C+ or below	0.201 (0.004)	0.182 (0.013)	0.188 (0.013)	0.136 (0.004)	0.161 (0.023)	0.133 (0.018)
B- to B+	0.492 (0.005)	0.501 (0.015)	0.505 (0.015)	0.443 (0.005)	0.443 (0.025)	0.436 (0.023)
A- or A	0.307 (0.005)	0.317 (0.013)	0.306 (0.013)	0.421 (0.005)	0.397 (0.022)	0.431 (0.022)
Plans to graduate from a 4-year college						
No	0.220 (0.005)	0.225 (0.013)	0.222 (0.013)	0.220 (0.004)	0.209 (0.024)	0.190 (0.020)
Yes	0.780 (0.005)	0.775 (0.013)	0.778 (0.013)	0.780 (0.004)	0.791 (0.024)	0.810 (0.020)
Truancy	1.733 (0.014)	1.706 (0.042)	1.723 (0.042)	1.632 (0.013)	1.719 (0.080)	1.674 (0.071)
Past 30-day alcohol use	2.065 (0.015)	2.123 (0.044)	2.088 (0.041)	1.552 (0.011)	1.452 (0.052)	1.479 (0.053)
Past 30-day cigarette use	1.698 (0.013)	1.703 (0.037)	1.689 (0.035)	1.098 (0.005)	1.101 (0.037)	1.094 (0.033)
Drug use sampling indicator						
Non-user	0.876 (0.004)	0.867 (0.008)	0.869 (0.008)	0.909 (0.003)	0.906 (0.014)	0.910 (0.012)
User	0.124 (0.004)	0.133 (0.008)	0.131 (0.008)	0.091 (0.003)	0.094 (0.014)	0.090 (0.012)
Metropolitan Status						
Large metropolitan area	0.304 (0.005)	0.282 (0.014)	0.297 (0.014)	0.335 (0.005)	0.332 (0.022)	0.346 (0.022)

Other metropolitan area	0.448 (0.005)	0.466 (0.015)	0.457 (0.015)	0.476 (0.005)	0.481 (0.024)	0.482 (0.023)
Non-metropolitan area	0.248 (0.005)	0.253 (0.013)	0.246 (0.012)	0.190 (0.004)	0.187 (0.022)	0.172 (0.018)

Notes: For the 2000 sample: *n* (12th grade) = 13,286; *n* (panel cohort participating at follow-up 1) = 1,429. For the 2019 sample: *n* (12th grade) = 13,713; *n* (panel cohort participating at follow-up 1) = 725.

Table 5. Comparison of cohort-specific multistep and pooled panel analysis weighting approaches in estimates of past 30-day alcohol and marijuana use prevalence at follow-up 1 (age 19/20): 12th grade panel cohorts of 2000 and 2019

Panel cohort		Proportion	(95% CI)
	Past 30-day alcohol use		
2000			
	Cohort-specific $PAW_{multistep}$	0.591	(0.561, 0.620)
	Cohort-specific PAW_{pooled}	0.589	(0.559, 0.618)
2019			
	Cohort-specific $PAW_{multistep}$	0.431	(0.385, 0.478)
	Cohort-specific PAW_{pooled}	0.430	(0.385, 0.474)
	Past 30-day marijuana use		
2000			
	Cohort-specific $PAW_{multistep}$	0.205	(0.183, 0.228)
	Cohort-specific PAW_{pooled}	0.206	(0.184, 0.229)
2019			
	Cohort-specific $PAW_{multistep}$	0.262	(0.215, 0.309)
	Cohort-specific PAW_{pooled}	0.259	(0.217, 0.302)

Notes: Unweighted sample sizes (*n*) for 2000 cohort = 1,396 for alcohol use and 1,415 for marijuana use; for 2019 cohort, 695 for alcohol use and 697 for marijuana use.

Table 6. Descriptive estimates of cohort-specific final pooled panel analysis weights (PAW_{pooled}) averaged across all panel cohorts from 1976–2020 as of calendar year 2021 data collection

Cohort-specific PAW_{pooled}		**N**	**Mean**	**(SD)**	**Min**	**Max**
PAW1	Cohort-specific Panel Analysis Weight at age 19/20 (FU1)	66,082	10.5	(6.2)	0.7	191.2
PAW2	Cohort-specific Panel Analysis Weight at age 21/22 (FU2)	60,321	11.2	(6.8)	0.9	276.3
PAW3	Cohort-specific Panel Analysis Weight at age 23/24 (FU3)	55,194	11.7	(6.9)	0.9	122.9
PAW4	Cohort-specific Panel Analysis Weight at age 25/26 (FU4)	49,802	12.4	(7.6)	1.2	141.3
PAW5	Cohort-specific Panel Analysis Weight at age 27/28 (FU5)	45,562	13.0	(8.3)	1.1	163.0
PAW6	Cohort-specific Panel Analysis Weight at age 29/30 (FU6)	41,773	13.5	(8.6)	1.2	162.2
PAW35	Cohort-specific Panel Analysis Weight at age 35 (FZ1)	32,924	14.4	(9.0)	1.3	158.4
PAW40	Cohort-specific Panel Analysis Weight at age 40 (FZ2)	26,613	15.0	(9.1)	1.5	155.9
PAW45	Cohort-specific Panel Analysis Weight at age 45 (FZ3)	20,963	15.3	(9.4)	1.8	129.7
PAW50	Cohort-specific Panel Analysis Weight at age 50 (FZ4)	15,621	15.3	(9.2)	2.4	106.0
PAW55	Cohort-specific Panel Analysis Weight at age 55 (FZ5)	10,255	15.1	(8.5)	2.4	79.0
PAW60	Cohort-specific Panel Analysis Weight at age 60 (FZ6)	4,586	14.7	(8.2)	2.8	76.8

Notes: PAW=panel analysis weight. Weight values shown are for panel cohort participants responding at the noted follow-up wave.

Table 7A. Comparison of 12th grade sociodemographic characteristics: Full 1976–2020 12th grade samples and panel cohorts responding at follow-up 1 (age 19/20)

	All 12th Grade Students: 12th grade sampling weight		Panel cohorts responding at follow-up 1: Unweighted		Traditional panel sampling weight[a]		Cohort-specific PAW1	
	%	(SE)	%	(SE)	%	(SE)	%	(SE)
Sex								
Female	51.3	(0.1)	56.1	(0.2)	57.2	(0.2)	51.5	(0.2)
Male	48.7	(0.1)	43.9	(0.2)	42.8	(0.2)	48.5	(0.2)
Race/ethnicity								
Black	12.2	(0.0)	8.7	(0.1)	9.4	(0.1)	12.2	(0.2)
Hispanic	10.0	(0.0)	7.6	(0.1)	8.0	(0.1)	10.0	(0.2)
White	69.2	(0.1)	76.4	(0.2)	75.0	(0.2)	69.1	(0.2)
Other (Asian, multiracial, other)	8.6	(0.0)	7.4	(0.1)	7.5	(0.1)	8.7	(0.1)
Parental education								
No college degree	55.0	(0.1)	53.2	(0.2)	52.6	(0.2)	54.6	(0.2)
1+ parent college degree	45.0	(0.1)	46.8	(0.2)	47.4	(0.2)	45.4	(0.2)
Region (in 12th grade)								
Midwest	26.1	(0.1)	28.7	(0.2)	28.6	(0.2)	26.0	(0.2)
Northeast	20.1	(0.1)	20.9	(0.2)	20.5	(0.2)	20.3	(0.2)
South	34.6	(0.1)	31.9	(0.2)	32.5	(0.2)	34.5	(0.2)
West	19.2	(0.1)	18.5	(0.2)	18.4	(0.2)	19.2	(0.2)

Notes: Total possible *n* for all 12th grade students = 688,290. Total possible *n* for those selected for follow-up and responding at follow-up 1 (age 19/20) = 66,082. Cohort-specific PAW1 = final cohort-specific pooled panel analysis weight for follow-up wave 1 (age 19/20), not smoothed or trimmed. Only the first half-sample of the 2020 panel cohort had the opportunity to respond at first follow-up as of 2021 data collection.

[a]The traditional panel sampling weight historically provided with MTF datasets that addresses only panel sample selection probabilities.

Table 7B. Comparison of 12th grade sociodemographic characteristics: Full 1976–2004 12th grade samples and panel cohorts responding at age 35

	All 12th Grade Students: 12th grade sampling weight		Panel cohorts responding at age 35: Unweighted		Traditional panel sampling weight[a]		Cohort-specific PAW35	
	%	(SE)	%	(SE)	%	(SE)	%	(SE)
Sex								
Female	51.2	(0.1)	57.4	(0.3)	58.0	(0.3)	51.6	(0.3)
Male	48.8	(0.1)	42.6	(0.3)	42.0	(0.3)	48.4	(0.3)
Race/ethnicity								
Black	12.2	(0.1)	6.4	(0.1)	7.1	(0.2)	12.1	(0.3)
Hispanic	6.9	(0.0)	4.4	(0.1)	4.7	(0.1)	7.0	(0.2)
White	74.0	(0.1)	84	(0.2)	82.9	(0.2)	74.1	(0.3)
Other (Asian, multiracial, other)	6.9	(0.0)	5.2	(0.1)	5.3	(0.1)	6.9	(0.2)
Parental education								
No college degree	57.8	(0.1)	55.3	(0.3)	54.9	(0.3)	57.7	(0.3)
1+ parent college degree	42.2	(0.1)	44.7	(0.3)	45.1	(0.3)	42.3	(0.3)
Region (in 12th grade)								
Midwest	27.5	(0.1)	32.1	(0.3)	32.0	(0.3)	27.4	(0.3)
Northeast	21.2	(0.1)	21.8	(0.2)	21.4	(0.2)	21.4	(0.3)
South	33.4	(0.1)	29.2	(0.3)	29.8	(0.3)	33.4	(0.3)
West	17.9	(0.1)	16.9	(0.2)	16.8	(0.2)	17.8	(0.3)

Notes: Total possible *n* for all 12th grade students = 467,435. Total possible *n* for those selected for follow-up and responding at age 35 = 32,924. Cohort-specific PAW35 = final cohort-specific pooled panel analysis weight for age 35 data collection wave, not smoothed or trimmed.

[a]The traditional panel sampling weight historically provided with MTF datasets that addresses only panel sample selection probabilities.

Table 8a. Substance use prevalence estimates by weighting approach at follow-up 1 (age 19/20): 12th grade cohorts 1976–2020

	Unweighted		Traditional panel sampling weight[a]		Cohort-specific PAW1		Unweighted		Traditional panel sampling weight		Cohort-specific PAW1	
	%	(SE)	%	(SE)	%	(SE)	%	(SE)	%	(SE)	%	(SE)
	Alcohol use (past 30 days)						**Cigarette use (past 30 days)**					
Sex												
Female	62.3	(0.3)	59.0	(0.3)	58.0	(0.3)	30.6	(0.2)	25.5	(0.2)	26.4	(0.3)
Male	69.9	(0.3)	65.8	(0.3)	64.2	(0.3)	31.5	(0.3)	26.0	(0.3)	27.4	(0.3)
Race/ethnicity												
Black	44.5	(0.7)	41.9	(0.7)	41.3	(0.8)	18.8	(0.5)	16.0	(0.5)	16.2	(0.5)
Hispanic	51.4	(0.7)	47.9	(0.7)	48.4	(0.8)	19.3	(0.6)	15.6	(0.5)	16.8	(0.6)
White	70.4	(0.2)	66.9	(0.2)	67.3	(0.2)	34.1	(0.2)	28.5	(0.2)	30.8	(0.2)
Other	54.8	(0.7)	50.7	(0.7)	51.6	(0.9)	25.2	(0.6)	20.3	(0.6)	22.1	(0.7)
Parental education												
No college degree	63.7	(0.3)	59.6	(0.3)	58.7	(0.3)	32.6	(0.3)	27.1	(0.2)	28.2	(0.3)
1+ parent college degree	68.3	(0.3)	65.0	(0.3)	64.4	(0.3)	29.2	(0.3)	24.1	(0.3)	25.3	(0.3)
	Marijuana use (past 30 days)						**Drugs other than marijuana (past year)**					
Sex												
Female	23.0	(0.2)	18.0	(0.2)	19.4	(0.2)	24.5	(0.2)	17.6	(0.2)	18.4	(0.2)
Male	30.8	(0.3)	23.2	(0.3)	25.1	(0.3)	29.1	(0.3)	20.4	(0.2)	21.4	(0.3)
Race/ethnicity												
Black	20.1	(0.5)	16.8	(0.5)	18.2	(0.6)	11.1	(0.4)	8.4	(0.4)	8.7	(0.4)
Hispanic	20.2	(0.6)	15.7	(0.5)	18.5	(0.7)	18.9	(0.6)	13.5	(0.5)	14.7	(0.6)
White	28.0	(0.2)	21.3	(0.2)	23.5	(0.2)	29.4	(0.2)	21.0	(0.2)	22.9	(0.2)
Other	23.4	(0.6)	17.6	(0.6)	19.9	(0.7)	21.4	(0.6)	14.8	(0.5)	16.4	(0.6)
Parental education												
No college degree	25.3	(0.2)	19.0	(0.2)	20.9	(0.3)	26.0	(0.2)	18.1	(0.2)	19.0	(0.2)
1+ parent college degree	27.6	(0.3)	21.6	(0.2)	23.6	(0.3)	27.3	(0.3)	19.7	(0.2)	21.2	(0.3)

Notes: Total possible *n* for those selected for follow-up and responding at follow-up 1 (age 19/20) = 66,082. Cohort-specific PAW1 = final cohort-specific pooled panel analysis weight for follow-up wave 1 (age 19/20), not smoothed or trimmed. Only the first half-sample of the 2020 panel cohort had the opportunity to respond at first follow-up as of 2021 data collection.
[a]The traditional panel sampling weight historically provided with MTF datasets that addresses only panel sample selection probabilities.

Table 8b. Substance use prevalence estimates by weighting approach at age 35: 12th grade cohorts 1976–2004

	Unweighted		Traditional panel sampling weight[a]		Cohort-specific PAW35		Unweighted		Traditional panel sampling weight		Cohort-specific PAW35	
	%	(SE)	%	(SE)	%	(SE)	%	(SE)	%	(SE)	%	(SE)
	Alcohol use (past 30 days)						**Cigarette use (past 30 days)**					
Sex												
Female	63.9	(0.4)	62.5	(0.4)	61.9	(0.4)	21.1	(0.3)	17.6	(0.3)	19.7	(0.3)
Male	75.9	(0.4)	74.7	(0.4)	73.8	(0.4)	24.7	(0.4)	20.4	(0.4)	22.9	(0.4)
Race/ethnicity												
Black	54.9	(1.1)	53.5	(1.1)	56.0	(1.1)	18.7	(0.9)	16.9	(0.9)	18.2	(0.9)
Hispanic	65.1	(1.3)	64.1	(1.3)	66.5	(1.3)	13.8	(0.9)	11.7	(0.9)	13.0	(0.9)
White	70.6	(0.3)	69.4	(0.3)	70.0	(0.3)	23.3	(0.3)	19.2	(0.2)	22.3	(0.3)
Other	62.9	(1.2)	61.5	(1.2)	62.3	(1.2)	22.8	(1.1)	18.7	(1.0)	21.7	(1.0)
Parental education												
No college degree	65.2	(0.4)	63.6	(0.4)	63.9	(0.4)	25.0	(0.3)	20.9	(0.3)	23.1	(0.3)
1+ parent college degree	74.1	(0.4)	73.1	(0.4)	73.3	(0.4)	19.4	(0.3)	16.0	(0.3)	18.5	(0.3)
	Marijuana use (past 30 days)						**Drugs other than marijuana (past year)**					
Sex												
Female	8.8	(0.2)	6.7	(0.2)	7.6	(0.2)	12.8	(0.3)	10.4	(0.2)	11.4	(0.3)
Male	15.8	(0.3)	11.7	(0.3)	13.3	(0.3)	16.2	(0.3)	12.4	(0.3)	14.4	(0.3)
Race/ethnicity												
Black	9.7	(0.7)	8.3	(0.6)	9.6	(0.7)	8.5	(0.7)	7.2	(0.6)	8.0	(0.7)
Hispanic	9.0	(0.8)	6.8	(0.7)	7.9	(0.7)	14.5	(1.0)	11.3	(0.9)	12.7	(1.0)
White	12.0	(0.2)	8.9	(0.2)	10.7	(0.2)	14.7	(0.2)	11.6	(0.2)	13.6	(0.2)
Other	12.0	(0.8)	8.8	(0.7)	10.1	(0.7)	14.0	(0.9)	11.7	(0.9)	13.2	(0.9)
Parental education												
No college degree	11.3	(0.2)	8.2	(0.2)	9.7	(0.2)	14.1	(0.3)	10.9	(0.3)	12.4	(0.3)
1+ parent college degree	12.4	(0.3)	9.5	(0.3)	11.3	(0.3)	14.5	(0.3)	11.8	(0.3)	13.5	(0.3)

Notes: Total possible *n* for those selected for follow-up and responding at age 35 = 32,924. Cohort-specific PAW35 = final cohort-specific pooled panel analysis weight for follow-up age 35, not smoothed or trimmed.

[a]The traditional panel sampling weight historically provided with MTF datasets that addresses only panel sample selection probabilities.

Table 9. Descriptive estimates of age-specific final pooled panel analysis weights (PAW_{pooled}) averaged across all panel cohorts from 1976-2020 as of calendar year 2021 data collection

Age-specific PAW_{pooled}		N	Mean	(SD)	Min	Max
PAW19	Age-specific Panel Analysis Weight at age 19	33,928	20.3	(11.7)	0.9	249.0
PAW20	Age-specific Panel Analysis Weight at age 20	32,154	21.5	(13.1)	1.7	264.1
PAW21	Age-specific Panel Analysis Weight at age 21	30,999	21.9	(12.8)	2.0	335.1
PAW22	Age-specific Panel Analysis Weight at age 22	29,321	22.6	(14.3)	2.2	250.2
PAW23	Age-specific Panel Analysis Weight at age 23	28,246	23.0	(13.5)	2.4	311.0
PAW24	Age-specific Panel Analysis Weight at age 24	26,947	23.6	(14.6)	2.2	233.3
PAW25	Age-specific Panel Analysis Weight at age 25	25,361	24.5	(14.8)	2.5	247.7
PAW26	Age-specific Panel Analysis Weight at age 26	24,441	24.9	(15.8)	2.1	303.2
PAW27	Age-specific Panel Analysis Weight at age 27	23,340	25.5	(15.7)	2.1	286.8
PAW28	Age-specific Panel Analysis Weight at age 28	22,221	26.2	(17.3)	2.1	311.7
PAW29	Age-specific Panel Analysis Weight at age 29	21,316	26.5	(16.7)	2.7	289.3
PAW30	Age-specific Panel Analysis Weight at age 30	20,456	26.8	(17.0)	2.0	261.0
PAW35	Age-specific Panel Analysis Weight at age 35	32,924	14.4	(9.0)	1.3	158.4
PAW40	Age-specific Panel Analysis Weight at age 40	26,613	15.0	(9.1)	1.5	155.9
PAW45	Age-specific Panel Analysis Weight at age 45	20,963	15.3	(9.4)	1.8	129.7
PAW50	Age-specific Panel Analysis Weight at age 50	15,621	15.3	(9.2)	2.4	106.0
PAW55	Age-specific Panel Analysis Weight at age 55	10,255	15.1	(8.5)	2.4	79.0
PAW60	Age-specific Panel Analysis Weight at age 60	4,586	14.7	(8.2)	2.8	76.8

Notes: PAW=age-specific panel analysis weight. Weight values shown are for panel cohort participants responding at the noted age.

Supplement Table 1. Monitoring the Future Panel Study administration timeline through calendar year 2021 data collection

12th grade	FU1: Age 19/20		FU2: Age 21/22		FU3: Age 23/24		FU4: Age 25/26		FU5: Age 27/28		FU6: Age 29/30		FZ1: Age 35	FZ2: Age 40	FZ3: Age 45	FZ4: Age 50	FZ5: Age 55	FZ6: Age 60
1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1993	1998	2003	2008	2013	2018
1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1994	1999	2004	2009	2014	2019
1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1995	2000	2005	2010	2015	2020
1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1996	2001	2006	2011	2016	2021
1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1997	2002	2007	2012	2017	
1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1998	2003	2008	2013	2018	
1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1999	2004	2009	2014	2019	
1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	2000	2005	2010	2015	2020	
1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	2001	2006	2011	2016	2021	
1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	2002	2007	2012	2017		
1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	2003	2008	2013	2018		
1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2004	2009	2014	2019		
1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2005	2010	2015	2020		
1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2006	2011	2016	2021		
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2007	2012	2017			
1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2008	2013	2018			
1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2009	2014	2019			
1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2010	2015	2020			
1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2011	2016	2021			
1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2012	2017				
1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2013	2018				
1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2014	2019				
1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2015	2020				
1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2016	2021				
2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2017					
2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2018					
2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2019					
2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2020					
2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2021					
2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017						
2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018						
2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019						
2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020						
2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021						
2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021[a]							
2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021								
2012	2013	2014	2015	2016	2017	2018	2019	2020	2021[a]									
2013	2014	2015	2016	2017	2018	2019	2020	2021										
2014	2015	2016	2017	2018	2019	2020	2021[a]											
2015	2016	2017	2018	2019	2020	2021												
2016	2017	2018	2019	2020	2021[a]													
2017	2018	2019	2020	2021														
2018	2019	2020	2021[a]															
2019	2020	2021																
2020	2021[a]																	
2021																		

[a]Only the first half-sample of the noted cohort will have completed panel participation at the specified follow-up age in the 2021 calendar year.

Supplement Table 2. Monitoring the Future Panel Study response rates (in proportions) by cohort as of calendar year 2021 data collection

Cohort	Total *n* selected per cohort	Follow-up and modal age: FU1: 19/20	FU2: 21/22	FU3: 23/24	FU4: 25/26	FU5: 27/28	FU6: 29/30	FZ1: 35	FZ2: 40	FZ3: 45	FZ4: 50	FZ5: 55	FZ6: 60
1976	2224	0.699	0.787	0.780	0.755	0.728	0.723	0.656	0.622	0.587	0.561	0.533	0.516
1977	2358	0.836	0.813	0.802	0.754	0.744	0.708	0.671	0.620	0.613	0.575	0.544	0.528
1978	2411	0.837	0.815	0.784	0.746	0.726	0.692	0.640	0.604	0.590	0.566	0.527	0.515
1979	2437	0.838	0.800	0.748	0.753	0.714	0.687	0.638	0.578	0.571	0.546	0.524	0.511
1980	2458	0.842	0.791	0.763	0.747	0.704	0.681	0.634	0.579	0.554	0.544	0.520	
1981	2458	0.836	0.788	0.773	0.726	0.697	0.674	0.598	0.562	0.538	0.505	0.478	
1982	2437	0.825	0.769	0.739	0.693	0.663	0.617	0.567	0.563	0.518	0.490	0.448	
1983	2426	0.798	0.778	0.722	0.684	0.646	0.594	0.541	0.538	0.492	0.461	0.453	
1984	2438	0.785	0.752	0.703	0.676	0.619	0.579	0.493	0.484	0.454	0.438	0.429	
1985	2467	0.781	0.713	0.679	0.631	0.578	0.542	0.474	0.454	0.449	0.414		
1986	2461	0.747	0.704	0.661	0.603	0.556	0.517	0.485	0.454	0.414	0.379		
1987	2466	0.724	0.686	0.646	0.582	0.544	0.525	0.486	0.441	0.404	0.370		
1988	2456	0.756	0.722	0.656	0.601	0.584	0.537	0.502	0.458	0.428	0.428		
1989	2478	0.730	0.690	0.622	0.565	0.528	0.506	0.456	0.411	0.389	0.383		
1990	2470	0.720	0.645	0.586	0.537	0.515	0.478	0.417	0.397	0.365			
1991	2473	0.736	0.658	0.603	0.541	0.510	0.482	0.429	0.387	0.362			
1992	2479	0.741	0.670	0.627	0.562	0.529	0.519	0.452	0.416	0.370			
1993	2449	0.689	0.647	0.588	0.534	0.495	0.499	0.426	0.375	0.351			
1994	2467	0.680	0.640	0.569	0.520	0.508	0.469	0.411	0.367	0.362			
1995	2469	0.677	0.618	0.565	0.523	0.530	0.482	0.426	0.394				
1996	2440	0.671	0.600	0.541	0.525	0.500	0.473	0.419	0.369				
1997	2457	0.627	0.559	0.520	0.523	0.468	0.457	0.383	0.341				
1998	2454	0.622	0.542	0.531	0.487	0.444	0.437	0.364	0.332				
1999	2459	0.592	0.514	0.544	0.465	0.450	0.436	0.367	0.353				
2000	2456	0.583	0.562	0.524	0.467	0.450	0.424	0.358					
2001	2448	0.539	0.555	0.482	0.452	0.442	0.409	0.340					
2002	2453	0.591	0.554	0.503	0.478	0.459	0.425	0.360					
2003	2449	0.588	0.510	0.473	0.448	0.406	0.384	0.346					
2004	2450	0.564	0.503	0.479	0.438	0.399	0.364	0.331					
2005	2450	0.517	0.507	0.475	0.410	0.386	0.366						
2006	2452	0.488	0.484	0.454	0.421	0.371	0.376						
2007	2452	0.536	0.478	0.462	0.410	0.383	0.381						
2008	2454	0.534	0.479	0.435	0.384	0.386	0.380						
2009	2453	0.491	0.441	0.389	0.331	0.335	0.335						
2010	2450	0.467	0.413	0.378	0.345	0.349	0.327						
2011	2450	0.454	0.391	0.348	0.332	0.346							
2012	2452	0.423	0.366	0.364	0.369	0.335							
2013	2450	0.402	0.344	0.345	0.373								
2014	2450	0.365	0.349	0.360	0.373								
2015	2451	0.322	0.325	0.363									
2016[a]	2450	0.333	0.347	0.341									
2017[a]	2452	0.340	0.333										
2018[a]	2450	0.348	0.345										
2019	2447	0.369											
2020[b]	1225[e]	0.440											
2021													

Notes: Response rates provided are unweighted. Red font indicates response rates including only half of the selected cohort, as the other half is not yet eligible for the relevant follow-up wave.

Supplement Appendix A: Probability of Panel Sample Eligibility

Analyses that were conducted to test and refine logistic regression models calculating individual probabilities of panel sample eligibility (i.e., providing both contact information and non-missing data on sex at the 12th grade survey) involved three steps: 1) selection of covariates, 2) comparison of the degree to which covariates retained explanatory associations across cohorts, and 3) year-specific modeling of the individual probability of providing both contact information and data on sex.

The Monitoring the Future (MTF) study collected data from annual nationally representative (i.e., national probability) samples of 12th grade students from 1976 through 2020 in the spring of each year. Detailed methods are described elsewhere (Miech et al., 2021a, 2022; Johnston et al., 2022). In brief, data collection for the 1976–2018 samples involved in-school paper surveys from approximately 15,500 students per year (range 12,600–18,924); student response rates averaged 83% (range 77%–86%). Data collection for the 2019 sample involved in-school randomization to either paper or electronic tablets (Miech et al., 2021a); 13,713 students responded for an 80% response rate. For the 2020 sample, data collection involved in-school electronic tablets only and was halted on March 15, 2020 because of the COVID-19 pandemic. The resulting 2020 sample size (3,770) was approximately one-quarter the size of a typical data collection, with a 79% student response rate (Miech et al., 2021b, 2022). In total, data from 683,417 12th grade students were utilized in analyses examining panel sample eligibility.

The following covariates for predicting panel sample eligibility were selected for all 12th grade samples based on extensive past research on factors associated with concurrent substance use likelihood (Miech et al., 2022; Johnston et al., 2020) and exploratory models predicting provision of contact information: sex, race/ethnicity, parental education (used as a proxy for family of origin socioeconomic status), number of parents in the home,

average high school grades, truancy, college plans, substance use (alcohol, cigarettes, marijuana, and drugs other than marijuana), metropolitan status, and MTF sampling strata (a categorical measure ranging from 1–27).

For steps 1 and 2 (selection of covariates and comparison of the degree to which covariates retained explanatory associations across samples), missing data on covariates were addressed by conducting a single imputation (PROC MI with fully conditional specification) using combined data from all samples. Analyses using multiple imputation with 10 imputed datasets were also explored in models predicting whether participants provided both contact information and data on sex (i.e., eligibility for panel sample); synthesis of results was conducted using PROC MIANALYZE. No substantive differences in results were observed (results available from authors on request).

To examine the degree to which these covariates were significant predictors of the dichotomous outcome of panel sample eligibility across samples, "bins" of samples were coded as follows: 1976–1980, 1981–1985, 1986–1990; 1991–1995, 1996–2000, 2001–2005, 2006–2010; 2011–2015, 2016–2018. (Samples 2019–2021 were not included in these initial models due to changes in data collection methodology of paper, tablet, and web.) The PROC HPSPLIT procedure then was used to produce classification trees by bin, with grow = entropy and prune = cost/complexity. No post-hoc pruning procedures were used. Across bins, sensitivity ranged from 0.9645 to 1.000, specificity from 0.0021 to 0.1451, and AUC (area under the curve) from 0.5376 to 0.6423. All covariates were included in one or more bins.

Based on the classification tree results, it was decided to include all selected covariate main effects in sample-specific models (i.e., including only one 12^{th} grade sample or year of data collection per model) predicting individual probabilities of panel sample eligibility. By utilizing sample-specific models, analyses avoided the complications of modeling paper-only (samples 1976–2018), randomized paper and tablet (sample 2019), tablet only (sample 2020), and web only procedures (sample 2021). For these models, the

imputation process was conducted separately by sample using a single imputation. For all samples, PROC LOGISTIC models regressed the dichotomous outcome of panel sample eligibility on sex, race/ethnicity, parental education, number of parents in the home, average high school grades, truancy, college plans, alcohol use, cigarette use, marijuana use, use of drugs other than marijuana, metropolitan status, and MTF sampling strata; sampling weights were not used. For the 2019 sample, an additional dichotomous covariate was included indicating whether the participant was randomized to paper or tablet procedures. Predicted probabilities were requested using a SAS option in the LOGISTIC procedure. Results by sample are shown in Supplement Appendix Table A1. Mean individual probabilities of eligibility by sample ranged from 0.434 [SD 0.193] in 2020 to 0.933 [SD 0.039] in 1985, with the lowest individual probability in 2002 (value 0.010) and the highest individual probability in 1982 (value 0.994).

Supplement Appendix Table A1. Individual predicted probabilities of panel eligibility by 12th grade sample year, 1976–2020

		IP_1 Probability of eligible			
Sample	**N**	**Mean**	**SD**	**Min**	**Max**
1976	16599	0.918	0.046	0.548	0.983
1977	18238	0.919	0.053	0.405	0.986
1978	18924	0.904	0.050	0.464	0.975
1979	16661	0.924	0.048	0.473	0.983
1980	16524	0.922	0.051	0.376	0.985
1981	17999	0.926	0.047	0.574	0.989
1982	18348	0.919	0.055	0.370	0.994
1983	16947	0.929	0.048	0.466	0.989
1984	16499	0.924	0.049	0.437	0.993
1985	16502	0.933	0.039	0.586	0.988
1986	15713	0.926	0.054	0.342	0.983
1987	16843	0.930	0.043	0.511	0.985
1988	16795	0.926	0.046	0.535	0.989
1989	17142	0.929	0.054	0.338	0.985
1990	15676	0.924	0.051	0.537	0.988
1991	15483	0.931	0.046	0.400	0.989
1992	16251	0.921	0.049	0.401	0.991
1993	16763	0.914	0.049	0.389	0.983
1994	15929	0.904	0.054	0.446	0.980
1995	15876	0.918	0.045	0.566	0.978
1996	14823	0.900	0.057	0.427	0.975
1997	15963	0.891	0.068	0.258	0.978
1998	15780	0.860	0.083	0.385	0.979
1999	14056	0.851	0.083	0.368	0.975
2000	13286	0.817	0.117	0.193	0.985
2001	13304	0.801	0.121	0.114	0.972
2002	13544	0.797	0.114	0.010	0.949
2003	15200	0.810	0.100	0.212	0.963
2004	15222	0.789	0.124	0.081	0.967
2005	15378	0.798	0.117	0.178	0.972
2006	14814	0.784	0.107	0.236	0.946
2007	15132	0.805	0.085	0.309	0.971
2008	14577	0.788	0.105	0.156	0.957
2009	14268	0.781	0.098	0.199	0.951
2010	15127	0.759	0.103	0.229	0.950
2011	14855	0.768	0.114	0.225	0.961
2012	14343	0.760	0.112	0.242	0.966
2013	13180	0.767	0.084	0.387	0.949
2014	13015	0.718	0.109	0.203	0.940
2015	13730	0.693	0.118	0.149	0.924
2016	12600	0.741	0.122	0.191	0.937
2017	13522	0.740	0.111	0.286	0.978
2018	14502	0.714	0.142	0.171	0.953
2019	13713	0.607	0.139	0.088	0.874
2020	3770	0.435	0.193	0.030	0.825

Notes: Participants were considered to be eligible for inclusion in the panel based on the following sample-specific conditions: 1976–2015 provided contact information and responded as either female or male on sex; 2016–2018 provided contact information only; 2019–2020 provided contact information and responded as either female or male on sex.

Supplement Appendix B: Recent Changes to the MTF 12th Grade Sampling Weight

Recent changes have been made to the Monitoring the Future (MTF) 12th grade sampling weight in regard to how school-level nonresponse is addressed. For the 1976–2018 12th grade samples, MTF sampling statisticians tracked each "active" 12th grade school and identified "empty slots" and "partially empty slots" where schools failed to participate. School-level nonresponse was then accounted for by adjusting the 12th grade sampling weights for similar schools or small groups of similar schools. Outliers in weight value or influence (percentage of weighted *n* within the sample) were examined and could be trimmed, although trimming was done very rarely. If the only reasonable school-level nonresponse adjustment would create an outlier for one or more of the participating schools, a partial nonresponse adjustment would be made. For example, participating School A was the only reasonable match for nonresponding School B; however, up-weighting School A to account for itself and School B would create too large of a weight or would up-weight the *n* for School A too much. The decision, then, would be to up-weight School A to account for 50% or 75% of School B's anticipated weighted *n*, with the remainder of the nonresponse being unaccounted for. This approach would be a compromise between accounting for school-level nonresponse on an individual school basis, while simultaneously attempting to avoid creating outliers.

For the 2019 sample, "empty slots" were again identified, but instead of adjusting for school-level nonresponse by looking for similar participating schools to up-weight, nonresponse was adjusted for on a more aggregate level based on school public/private status, U.S. Census region (4 categories), and metropolitan status (3 categories). A total of 24 “cells” were formed based on the combination of school public/private status, region, and metropolitan status; some of these cells were collapsed (particularly within the private school sample) due to the initial cell(s) being too small. For example, through the 2018 sample, a participating school within the New York metropolitan area would be utilized in the nonresponse

adjustment for a non-participating school within the New York metropolitan area. For the 2019 sample, public school nonresponse within New York City was adjusted for throughout participating public schools in the relevant cell (including New York City, Boston, Philadelphia, etc.). All participating schools in the cell were up-weighted for all non-participating schools (empty slots) in the cell. Thus, nonresponse was spread more evenly, and individual schools were far less likely to be pushed to outlier values due to the nonresponse of similar schools.

Additional changes to the 12th grade sampling weight in regard to how school-level nonresponse is addressed were made for the 2020 and 2021 samples. While the analyses in the main text accompanying this supplemental appendix did not present data for samples after 2019, we briefly provided an overview of changes made for the 2020 and 2021 samples. Calculation of the 12th grade sampling weight for the 2020 sample no longer focused on the nonresponse ("empty slots") of individual schools; instead, the following procedure was used. First, probability weights were developed based on the number of cleaned surveys per school. Second, eight post-stratification cells (defined by the intersection of metropolitan/non-metropolitan area and region) were defined. School public/private status was not considered; public and private schools were mixed within the post-stratification cells. Third, each school's probability weight was post-stratified such that when weighted, each of the eight cells matched their corresponding proportion on the current MTF sampling frame. For example, using initial school probability weights, 12th grade metropolitan schools in the Northeast Region accounted for 11.6% of the weighted sample. However, the same group comprised 15.4% of the MTF sampling frame. Hence, all schools in that cell received a post-stratification adjustment greater than 1.0. In the final calculation of the 12th grade sampling weight, the 12th grade metropolitan schools (more precisely, the responding students within those participating schools) in the Northeast Region accounted for 15.4% of the weighted sample.

Calculation of the 12th grade sampling weight for the 2021 sample followed the 2020 sample methodology with the following exception. The 2020 data collection had to be halted mid-study because of the COVID-19 pandemic (Miech et al., 2022), resulting in a smaller-than-normal number of participating schools. For the 2021 sample, the larger number of participating schools (compared with the 2020 sample) allowed use of 12 post-stratification cells versus 8. For the 2021 sample, these cells were defined based the intersection of on the 3-level metropolitan status measure (large metropolitan area, other metropolitan area, non-metropolitan area) and region. The difference with respect to the 2020 sample is that for the 2021 sample, for example, a Boston school (large metropolitan area, Northeast) and a Hartford school (other metropolitan area, Northeast) were not in the same post-stratification cell as they would have been in for the 2020 sample (when no distinction was drawn between large metropolitan area and other metropolitan area).

In summary, the key change in 12th grade sampling weight development relates to how school-level nonresponse is addressed. Post-stratifying weights with respect to metropolitan status and Census region has replaced school-specific nonresponse adjustments. This change addresses nonresponse within a larger pool (cell). In addition to school-level nonresponse, the new approach also adjusts across post-stratification cells for 1) disparities in student-level nonresponse (largely absenteeism) and 2) any disparities that may exist in enrollment estimates versus actual enrollment. It should be noted that some rather large 12th grade sampling weight values have been allowed. When the school probability weight post-stratification adjustment is completed and modified to the student population rather than the school population, individual 12th grade sampling weight values can reach double digits. This typically occurs only for a very small number of students, as well as for students who come from schools with a relatively small number of participants. When the 12th grade sampling weight is summed back to the school level, is it unlikely that the school will have a disproportionate value related to the total weighted *n* for a given year.

Supplement Appendix C: MTF Panel Analysis Weight Selection and Normalization

Use of MTF Panel data typically involves either wide or skinny/long analytic datasets. Wide datasets are those in which each respondent has only one line of data; these datasets can be rectangular (all respondents have the opportunity to participate in a shared specific age or wave of panel data collection) or non-rectangular (respondents are not expected to participate in the same ages or waves of data collection). Skinny or long data structures are those in which respondents can have multiple lines of data, and may or may not have had the opportunity to participate in shared ages/waves of data collection. Selection of the appropriate panel analysis weight depends on the analytic dataset structure. The language used in the remainder of this appendix will continue to use the age-specific pooled weight naming conventions used in the main text (i.e., PAW19, PAW20,...PAW60). However, Table C1 provides the actual MTF dataset variable names and labels users will find in the restricted panel datasets.

Supplement Appendix Table C1. MTF restricted panel data variable numbers and labels for age-specific final pooled panel analysis weights (PAW_{pooled})

Age-specific PAW_{pooled}	Variable number	Variable label
PAW19	V1632	90053:FU PANL ANALYSIS WT AGE19
PAW20	V1633	90054:FU PANL ANALYSIS WT AGE20
PAW21	V2632	90055:FU PANL ANALYSIS WT AGE21
PAW22	V2633	90056:FU PANL ANALYSIS WT AGE22
PAW23	V3632	90057:FU PANL ANALYSIS WT AGE23
PAW24	V3633	90058:FU PANL ANALYSIS WT AGE24
PAW25	V4632	90059:FU PANL ANALYSIS WT AGE25
PAW26	V4633	90060:FU PANL ANALYSIS WT AGE26
PAW27	V5632	90061:FU PANL ANALYSIS WT AGE27
PAW28	V5633	90062:FU PANL ANALYSIS WT AGE28
PAW29	V6632	90063:FU PANL ANALYSIS WT AGE29
PAW30	V6633	90064:FU PANL ANALYSIS WT AGE30
PAW35	V35031	90028:FZ PANL ANALYSIS WT FZ35
PAW40	V40031	90029:FZ PANL ANALYSIS WT FZ40

PAW45	V45031	90030:FZ PANL ANALYSIS WT FZ45
PAW50	V50031	90031:FZ PANL ANALYSIS WT FZ50
PAW55	V55031	90032:FZ PANL ANALYSIS WT FZ55
PAW60	V60031	90033:FZ PANL ANALYSIS WT FZ60

Notes: PAW=panel analysis weight.

The MTF age-specific panel analysis weights have been developed to weight back to the 12th grade nationally representative (i.e., national probability) samples. Each typical full panel cohort has an approximate size of 2,450, and each cohort half-sample has an approximate size of 1,225. In contrast, each typical annual 12th grade sample has an approximate size of 12,000–19,000. (As noted in the main text, the 2020 and 2021 12th grade sample sizes were smaller due to factors associated with the COVID-19 pandemic.) Using the age-specific panel analysis weights without normalization in analyses results in a weighted n that reflects the number of 12th grade students in the original samples that are represented by the panel half-samples at each age, which is much larger than the actual half-sample sizes selected for follow-up. If a particular research question involves combining data from multiple cohorts or ages, it is advisable to normalize the panel analysis weights for the analytic dataset so that each half-sample contributes to the analyses in proportion to its unweighted n.

In this appendix, we discuss weight selection and normalization guidelines for both wide/rectangular and skinny/long analytic datasets using Table C2 as a visual guide for wave or age of administration.

Supplement Appendix Table C2. Monitoring the Future Panel Study administration timeline through calendar year 2021 data collection

12th grade	FU1: Age 19/20		FU2: Age 21/22		FU3: Age 23/24		FU4: Age 25/26		FU5: Age 27/28		FU6: Age 29/30		FZ1: Age 35	FZ2: Age 40	FZ3: Age 45	FZ4: Age 50	FZ5: Age 55	FZ6: Age 60
1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1993	1998	2003	2008	2013	2018
1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1994	1999	2004	2009	2014	2019
1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1995	2000	2005	2010	2015	2020
1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1996	2001	2006	2011	2016	2021
1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1997	2002	2007	2012	2017	
1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1998	2003	2008	2013	2018	
1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1999	2004	2009	2014	2019	
1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	2000	2005	2010	2015	2020	
1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	2001	2006	2011	2016	2021	
1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	2002	2007	2012	2017		
1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	2003	2008	2013	2018		
1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2004	2009	2014	2019		
1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2005	2010	2015	2020		
1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2006	2011	2016	2021		
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2007	2012	2017			
1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2008	2013	2018			
1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2009	2014	2019			
1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2010	2015	2020			
1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2011	2016	2021			
1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2012	2017				
1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2013	2018				
1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2014	2019				
1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2015	2020				
1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2016	2021				
2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2017					
2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2018					
2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2019					
2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2020					
2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2021					
2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017						
2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018						
2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019						
2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020						
2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021						
2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021[a]							
2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021								
2012	2013	2014	2015	2016	2017	2018	2019	2020	2021[a]									
2013	2014	2015	2016	2017	2018	2019	2020	2021										
2014	2015	2016	2017	2018	2019	2020	2021[a]											
2015	2016	2017	2018	2019	2020	2021												
2016	2017	2018	2019	2020	2021[a]													
2017	2018	2019	2020	2021														
2018	2019	2020	2021[a]															
2019	2020	2021																
2020	2021[a]																	
2021																		

Notes: For follow-ups (FU) 1-6, each 12th grade cohort is divided into first and second half-samples. The first half-sample begins panel study participation one year after 12th grade (modal age 19) while the second half-sample begins two after 12th grade (modal age 20). Respondents are then surveyed every two years through follow-up 6.

Wide/rectangular analytic datasets

Analytic datasets in which each respondent has a single record or line of data are known as wide or rectangular datasets. Specific research questions using wide/rectangular datasets may require that all respondents: 1) participated at a shared specific final wave or age; 2) had the opportunity to participate at a shared specific wave or age; or 3) participated during shared calendar years of data collection. For situations 1 and 2, the recommended age-specific panel analysis weight is the weight for each respondent's final age of participation. For situation 3, the recommended age-specific panel analysis weight is the weight for each respondent for the relevant age conducted during the specified calendar year. We describe each of these three situations below.

1. Rectangular analytic dataset with a shared final wave of participation; one line of data per respondent. Example: Analysis using data from ages 19 through 30, requiring a response at follow-up 6 (either modal age 29 or 30) using data from 1990–1992 12th grade cohorts (see Tables C2 and C3, yellow highlight). All respondents would be assigned their follow-up 6 weight for analysis, which would then be normalized to the analytic sample prior to modeling.

Supplement Appendix Table C3. Rectangular dataset with shared final wave (follow-up 6) of participation

V101 (Cohort)	**MTFID**	**FU6 PAW29:** FU PANL ANALYSIS WT AGE29	**FU6 PAW30:** FU PANL ANALYSIS WT AGE30	**NPAW**
1990	A	15.9505		1.0989
1990	C		8.4589	0.5828
1990	D	10.0820		0.6946
1990	E		5.0186	0.3458
1991	H	35.7957		2.4661
1991	J		12.7972	0.8816
1992	K	9.5772		0.6598
1992	L		10.5748	0.7285
1992	N	22.3808		1.5419

Notes: FU=follow-up. Respondents with MTFID A, D, H, K, and N were assigned to the first half-sample for their respective cohorts and were thus surveyed at ages 19, 21, 23, 25, 27, and 29. Respondents with MTFID C, E, J, and L were assigned to the second half-sample for their respective cohorts, and were thus surveyed at ages 20, 22, 24, 26, 28, and 30.

Using the data shown in Table C3, the normalization process would proceed as follows. First, obtain the mean of the follow-up 6 weight values (PAW29 and PAW30 combined) for the nine cases in the analytic sample (mean=14.51508). Then, divide each respondent's follow-up 6 weight value by the obtained mean. By doing so, the sum of the normalized weights for the analytic dataset would be equal (with rounding) to the unweighted *n* of the analytic sample (here summing to 9).

Example SAS syntax for conducting normalization of the FU6 weight in Table C3 is provided here:

```
/* ************************************************************** */
/* Create a single, normalized weight variable across respondents */
/*          with a shared final wave of participation             */
/*                 in a rectangular dataset                       */
/*                                                                */
/*         In this example, each respondent has a value           */
/*            for either PAW29 or PAW30, the                      */
/*     age-specific panel analysis weights at follow-up (FU) 6.   */
/*                 It is assumed the PAW# of                       */
/*  interest are already in your working dataset libname.dataset  */
/* ************************************************************** */

/* *************************************************** */
/* Obtain and output the unweighted mean of the panel    */
/* analysis weights you wish to use. In this example,    */
/* V35 identifies the half-sample (1=first half-sample,  */
/* 2=second half-sample), PAW29 is the panel analysis    */
/* weight for age 29 at FU6, PAW30 is the panel analysis */
/* weight for age 30 at FU6,and mPAWFU6 is the           */
/* unweighted mean across all cases PAW29 + PAW30.       */
/* *************************************************** */

DATA newdataset;
  SET libname.dataset;
  /* make a new variable that combines PAW29 and PAW30 */
  /* into a single FU6 weight                          */
  IF (V35=1) then PAWFU6=PAW29;
     ELSE IF (V35=2) then PAWFU6=PAW30;
RUN;

PROC MEANS DATA = newdataset;
  VAR PAWFU6;
```

```
  OUTPUT OUT = forwgtmean (DROP = _TYPE_ _FREQ_) mean=mPAWFU6;
RUN;

/* ********************************************************* */
/* Merge the obtained mean, mPAWFU6, onto each record in     */
/* your working dataset newdataset                           */
/*                                                           */
/* Normalize the weight by dividing each respondent's        */
/* PAWFU6 by mPAWFU6, creating NPAW. Use NPAW in analysis    */
/* ********************************************************* */

DATA newdataset_2;
  /* this statement adds mPAWFU6 to each case */
  IF _n_ = 1 THEN SET forwgtmean ;
  SET newdataset;
  /* normalize each individual's weight */
  NPAW = PAWFU6/mPAFU6;
RUN;
```

2. **Rectangular analytic datasets with opportunity to participate at shared specific wave or age, but without shared final wave/age of participation; one line of data per respondent**. Example: research question focusing on a binary outcome indicating any prevalence of a specific behavior across ages 50, 55, or 60 using data from 1976–1977 cohorts (see Tables C2 and C4, teal highlight). In this example, respondents would be required only to participate at least once across modal ages 50, 55, and 60. Any respondent from the 1976 or 1977 cohorts who participated at modal age 60 would be assigned the FZ60 weight for analysis. Respondents who participated at modal age 55 (but not 60) would be assigned the FZ55 weight for analysis (e.g., MTFID “A” in Table C4), while those who participated at modal age 50 (but not 55 or 60) would be assigned the FZ50 weight for analysis (MTFID “H” in Table C4). All assigned weights would be normalized across waves to the analytic sample prior to modeling (the normalized weights should sum to the analytic sample).

Supplement Appendix Table C4. Rectangular dataset without shared final wave/age of participation

V101 (Cohort)	MTFID	PAW50: FU PANL ANALYSIS WT FZ50	PAW55: FU PANL ANALYSIS WT FZ55	PAW60: FU PANL ANALYSIS WT FZ60	NPAW
1976	A	.	22.1483	.	1.8453
1976	C	4.7584	5.1783	4.8184	0.4014
1976	D	5.2056	.	5.2757	0.4395
1976	E	11.3962	11.8252	12.4461	1.0369
1977	G	5.2207	5.1058	5.4717	0.4559
1977	H	20.3957	.	.	1.6992
1977	I	9.9941	10.8556	11.1877	0.9321
1977	J	11.7918	12.8631	14.2790	1.1896

Example SAS syntax for conducting normalization of the assigned panel analysis weight in Table C4 is provided below:

```
/* ************************************************************** */
/* Create a single, normalized weight variable across respondents */
/*        without a shared final wave of participation            */
/*                 in a rectangular dataset                       */
/*                                                                */
/*  In the MTF FZ panel data, V50001, V55001, and V60001 contain  */
/*  year of administration (YOA) data for ages 50, 55, and 60,    */
/*  respectively. The YOA variables will have valid data if a     */
/*  respondent completed a survey at that age.                    */
/*                                                                */
/*  Use YOA variables to assign the latest wave of participation  */
/*  panel analysis weights to a single variable, PAW, that can    */
/*  then be used to normalize the weights for analysis.           */
/*                                                                */
/*  In this example, it is assumed the PAW## of interest are      */
/*  already included in your working dataset libname.dataset      */
/* ************************************************************** */

DATA newdata;
  SET libname.dataset;
  /* Determine the panel analysis weight to use based on final wave of */
  /*  participation. If YOA is valid at the latest wave, assign it to  */
  /*  new variable PAW; if not, check the next wave, etc.              */
  IF (V60001 GT .Z) THEN PAW = PAW60;
    ELSE IF (V55001 GT .Z) THEN PAW = PAW55;
    ELSE IF (V50001 GT .Z) THEN PAW = PAW50;
  RUN;

/* ******************************************************* */
/* Obtain and output the single, unweighted mean of the    */
```

```
/*  panel analysis weight(s) you wish to use. In this      */
/*  example, PAW is the panel analysis weight across ages, */
/*  mPAW is the unweighted mean across all cases for PAW.  */
/* ******************************************************* */

/* obtain mean of single weight variable */

PROC MEANS DATA=newdata;
  VAR PAW;
  OUTPUT OUT=forwgtmean (DROP= _TYPE_ _FREQ_ ) mean=mPAW;
RUN;

/* *************************************************** */
/* Merge the obtained mean, mPAW, onto each record in  */
/*   your working dataset newdata                      */
/*                                                     */
/* Normalize the weight by dividing each respondent's  */
/*   PAW by mPAW, creating NPAW. Use NPAW in analysis  */
/* *************************************************** */

DATA libname.newdata2;
  /* this statement adds mPAW to each case */
  IF _n_ = 1 THEN SET forwgtmean ;
  SET newdata ;
  /* normalize each individual's weight */
  NPAW = PAW/mPAW;
RUN;
```

3. Rectangular analytic dataset with shared calendar years of data collection, but not shared waves or ages of data collection; one line of data per respondent. In these situations, the waves/ages of response vary across respondents. For example, consider a new measure introduced in 2005 for all waves of panel data collection; any data collected prior to 2005 would not include the measure. Analyses are proposed to obtain the prevalence of the new measure across ages 19–30 during the years 2005-2006. This results in using data from 1993–2005 12th grade cohorts responding at ages 19 through 30 during years 2005 and 2006 (see Tables C2 and C5, green highlight). For this example, a respondent from the 1999 cohort who participated at age 25 in 2006 would be assigned the age 25 weight for analysis (e.g., MTFID “L” in Table C5); a respondent from the 2004 cohort who participated at age 19 in 2005 would be assigned the age 19 weight (e.g., MTFID “V”); and a respondent from the 2004 cohort who

participated at age 20 in 2006 would be assigned the age 20 weight for analysis (e.g., MTFID “W”).

As shown in Table C5, most respondents would likely have weights for data wave collections that occurred both before and after their 2005/2006 participation. However, in this example, the focus is on analysis of data collected in particular 2005/2006 calendar years, and each respondent provided data in only one of the specified years. Thus, the weight for the age surveyed during the specified calendar year would be used. The assigned weights would be normalized prior to modeling (the normalized weights should sum to the analytic sample).

Supplement Appendix Table C5. Time-invariant non-rectangular dataset without shared final wave/age of participation

			FU1		FU2		FU3		FU4		FU5		FU6		
V101	**V35**	**MTFID**	**PAW19**	**PAW20**	**PAW21**	**PAW22**	**PAW23**	**PAW24**	**PAW25**	**PAW26**	**PAW27**	**PAW28**	**PAW29**	**PAW30**	**NPAW**
1993	2	A	.	18.1629	.	18.2501	.	20.7744	.	28.2666	.	39.1903	.	35.6415	2.1587
1994	1	B	15.8564	.	.	.	.	.	31.6492	.	31.6366	.	30.1456	.	1.8259
1994	2	C	.	13.2599	.	16.504	.	14.6813	.	18.9544	.	24.0395	.	22.4012	1.3568
1995	1	D	11.61	.	12.9629	.	13.3691	.	15.9308	.	.	.	16.1476	.	0.978
1995	2	E	.	11.301	.	12.1432	.	13.1835	.	.	.	15.5406	.	16.7575	0.9413
1996	1	F	15.1337	.	.	.	.	.	22.1345	.	27.9584	.	.	.	1.6934
1996	2	G	.	9.8248	.	10.4104	.	.	.	.	.	13.2952	.	.	0.8053
1997	1	H	9.6608	.	10.6624	.	10.9272	.	13.4282	.	17.7887	.	12.6746	.	1.0774
1997	2	I	.	9.7482	.	11.0633	.	10.1397	.	13.8375	.	17.9435	.	13.8262	0.8381
1998	1	J	6.2188	.	6.9535	.	7.2357	.	8.8278	.	8.86517	.	7.9512	.	0.5347
1998	2	K	.	10.0105	.	10.2341	.	.	.	13.136	.	.	.	.	0.7956
1999	1	L	13.4257	.	14.2508	.	.	.	14.9885	.	.	.	17.4659	.	0.9078
1999	2	M	.	18.3246	.	21.0918	.	22.7119	.	27.2951	.	40.1924	.	36.9173	1.3756
2000	1	N	12.5112	.	14.2945	.	14.1087	.	.	.	.	.	.	.	0.8545
2000	2	O	.	18.3245	.	21.0918	.	22.7119	.	27.295	.	40.1923	.	36.917	1.3756
2001	1	P	13.5839	.	15.2968	.	15.0516	.	18.535	.	25.2865	.	23.5639	.	0.9117
2001	2	Q	.	17.2851	.	19.4863	.	19.6214	.	.	.	.	.	.	1.1803
2002	1	R	8.1322	.	8.5179	.	10.1596	.	10.7534	.	11.7381	.	12.1477	.	0.5159
2002	2	S	.	13.6452	.	13.9882	.	16.2874	.	19.318	.	.	.	.	0.8472
2003	1	T	15.332	.	19.4723	.	19.501	.	.	.	.	.	.	.	1.1794
2003	2	U	.	11.4126	.	11.7064	.	15.6475	.	16.1472	.	19.7911	.	19.3984	0.6912
2004	1	V	7.1746	.	.	.	.	.	.	.	.	.	.	.	0.4346
2004	2	W	.	5.3045	.	.	.	6.0501	.	6.2725	.	6.7806	.	9.4421	0.3213
2005	1	X	6.5975	.	6.5675	.	.	.	.	.	6.679	.	.	.	0.3996

Notes: V101=cohort; V35=half-sample indicator (1=first half-sample, 2=second half-sample). FU=follow-up.

Example SAS syntax for conducting normalization of the assigned panel analysis weight in Table C5 is provided below:

```
/* ************************************************************** */
/* Create a single, normalized weight variable across respondents */
/*          without a shared final wave of participation          */
/*          in a time-invariant, non-rectangular dataset          */
/*                                                                */
/*  In this example, V101(12th grade year of administration[YOA]  */
/*  and V35 (R surveyed in first or second half-sample) are       */
/*  needed to determine the correct PAW. It is assumed the PAW#   */
/*  of interest are already included in your working dataset      */
/*  libname.dataset                                               */
/*                                                                */
/*  Use V101 and V35 to assign the correct wave PAW# to a single  */
/*  variable, PAW. PAW will then be used to normalize the weights */
/*  for analysis, NPAW.                                           */
/* ************************************************************** */

DATA newdata;
  SET libname.dataset;
  /* Assign analysis weight to use based on specified wave of */
  /* participation as determined by V101 (12th grade year of  */
  /*administration [YOA]) and V35 (half-sample)               */
  /*for your analyses                                         */
   IF (V101=2005) AND (V35=1) THEN PAW=PAW19;
  ELSE IF (V101=2004) AND (V35=1) THEN PAW=PAW19;
  ELSE IF (V101=2004) AND (V35=2) THEN PAW=PAW20;
  ELSE IF (V101=2003) AND (V35=1) THEN PAW=PAW21;
  ELSE IF (V101=2003) AND (V35=2) THEN PAW=PAW20;
  ELSE IF (V101=2002) AND (V35=1) THEN PAW=PAW21;
  ELSE IF (V101=2002) AND (V35=2) THEN PAW=PAW22;
  ELSE IF (V101=2001) AND (V35=1) THEN PAW=PAW23;
  ELSE IF (V101=2001) AND (V35=2) THEN PAW=PAW22;
  ELSE IF (V101=2000) AND (V35=1) THEN PAW=PAW23;
  ELSE IF (V101=2000) AND (V35=2) THEN PAW=PAW24;
  ELSE IF (V101=1999) AND (V35=1) THEN PAW=PAW25;
  ELSE IF (V101=1999) AND (V35=2) THEN PAW=PAW24;
  ELSE IF (V101=1998) AND (V35=1) THEN PAW=PAW25;
  ELSE IF (V101=1998) AND (V35=2) THEN PAW=PAW26;
  ELSE IF (V101=1997) AND (V35=1) THEN PAW=PAW27;
  ELSE IF (V101=1997) AND (V35=2) THEN PAW=PAW26;
  ELSE IF (V101=1996) AND (V35=1) THEN PAW=PAW27;
  ELSE IF (V101=1996) AND (V35=2) THEN PAW=PAW28;
  ELSE IF (V101=1995) AND (V35=1) THEN PAW=PAW29;
  ELSE IF (V101=1995) AND (V35=2) THEN PAW=PAW28;
  ELSE IF (V101=1994) AND (V35=1) THEN PAW=PAW29;
  ELSE IF (V101=1994) AND (V35=2) THEN PAW=PAW30;
  ELSE IF (V101=1993) AND (V35=2) THEN PAW=PAW30;
RUN;
```

```
/* ************************************************** */
/* Obtain and output the single, unweighted mean of the  */
/*  panel analysis weight(s) you wish to use. In this    */
/*  example, PAW is the panel analysis weight for each   */
/*  V101/V35 combination that defines your sample, mPAW  */
/*  is the unweighted mean across all cases for PAW.     */
/* ************************************************** */

PROC MEANS DATA=newdata;
  VAR PAW;
  OUTPUT OUT=forwgtmean (DROP= _TYPE_ _FREQ_ ) mean=mPAW;
RUN;

/* ************************************************ */
/* Merge the obtained mean, mPAW, onto each record in  */
/*   your working dataset newdata                      */
/*                                                     */
/* Normalize the weight by dividing each respondent's  */
/*   PAW by mPAW, creating NPAW. Use NPAW in analysis  */
/* ************************************************ */

DATA libname.newdata2;
  /* this statement adds mPAW to each case */
  IF _n_ = 1 THEN SET forwgtmean ;
  SET newdata ;
  /* normalize each individual's weight */
  NPAW = PAW/mPAW;
RUN;
```

Skinny or long analytic datasets

In skinny/long datasets, respondents can have multiple lines of data; each line of data represents a different wave of participation. For analyses using such datasets, the recommended panel analysis weighting approach depends on whether or not the research question focuses on marginal (population-averaged) models or mixed (subject-specific) models focusing on multi-level modeling of the longitudinal repeated-measures analysis (and, by extension, whether the analytic software allows two-level weights).

1. Marginal (population-averaged) models. For these analyses, the goal is to fit a single model across respondents while controlling for within-person correlation. For such models, it is recommended that analysts use the panel analysis weight associated with each wave of participation. As an example, consider an analysis with an outcome measured from ages 35–60, where respondents have a minimum of 1 line of data (for those who participated

only at one wave between 35 and 60) and a maximum of six lines of data (for those who participated at ages 35, 40, 45, 50, 55, and 60.) Each line of data would be assigned the panel analysis weight for that wave. The assigned weights would be normalized across respondents and waves (summing to the analytic sample *n*) prior to modeling.

Table C6 provides an example for 4 hypothetical respondents, along with their respective wave-specific panel analysis weights. The first respondent (MTFID "A") participated only at FZ35, the second (MTFID "B") participated in all 6 waves from FZ35 to FZ60, the third (MTFID "C") participated in FZ55 and FZ60, and the last (MTFID "D") participated in FZ35, FZ50 and FZ60.

Supplement Appendix Table C6. Initial structure of hypothetical panel data - wide data set

V101 (Chort)	**MTFID**	**PAW35:** FU PANL ANALYSIS WT FZ35	**PAW40:** FU PANL ANALYSIS WT FZ40	**PAW45:** FU PANL ANALYSIS WT FZ45	**PAW50:** FU PANL ANALYSIS WT FZ50	**PAW55:** FU PANL ANALYSIS WT FZ55	**PAW60:** FU PANL ANALYSIS WT FZ60
1976	A	13.7705	.	.	.	.	.
1976	B	4.3059	4.2477	4.3142	4.4525	4.4221	4.4016
1977	C	.	.	.	.	12.225	11.966
1977	D	8.1337	.	.	9.7775	.	12.374

Table C7 provides the values for the skinny/long panel analysis weight assigned for analysis for each respondent/wave.

Supplement Appendix Table C7. Skinny/Long data structure for analysis

V101 (Cohort)	**MTFID**	**FZ**	**Skinny Panel Analysis Weight Prior to Normalization (PAW)**	**NPAW**
1976	A	35	13.7705	1.7507
1976	B	35	4.3059	0.5474
1976	B	40	4.2477	0.5400
1976	B	45	4.3142	0.5485
1976	B	50	4.4525	0.5661
1976	B	55	4.4221	0.5662
1976	B	60	4.4016	0.5596
1977	C	55	12.225	1.5542
1977	C	60	11.966	1.5213
1977	D	35	8.1337	1.0340
1977	D	50	9.7775	1.2430
1977	D	60	12.374	1.5731

Example SAS syntax for conducting normalization of the assigned panel analysis weight in Table C7 is provided below:

```
/* ************************************************************ */
/* Create a single, normalized weight variable across respondents */
/*              when using a long/skinny data file for            */
/*                  non-multilevel analysis models                */
/*                                                                */
/*  In this example, a "skinny" file is used, where each record   */
/*  represents a wave-specific response. The number of waves per  */
/*  respondent can vary, and each record contains a wave-specific */
/*  panel analysis weight PAW. It is assumed the PAW# are already */
/*  included in your working dataset libname.skinnydataset        */
/*                                                                */
/*  In a long/skinny dataset, there is a single variable for PAW. */
/*  PAW is used to create the mean, mPAW. These variables will    */
/*  then be used to normalize the weight for analysis, NPAW.      */
/* ************************************************************ */

/* *************************************************** */
/* Obtain and output the single, unweighted mean of the  */
/*  panel analysis weight you wish to use. In this       */
/*  example, PAW is the panel analysis weight for each   */
/*  row of data in your long/skinny dataset. mPAW is the */
/*  unweighted mean across all cases for PAW.            */
/* *************************************************** */

PROC MEANS DATA=libname.skinnydataset;
  VAR PAW;
  OUTPUT OUT=forwgtmean (DROP= _TYPE_ _FREQ_ ) mean=mPAW;
RUN;

/* ************************************************* */
/* Merge the obtained mean, mPAW, onto each record in  */
/*   your working dataset newdata                      */
/*                                                     */
/* Normalize the weight by dividing each respondent's  */
/*   PAW by mPAW, creating NPAW. Use NPAW in analysis  */
/* ************************************************* */

DATA libname.skinnydataset2
  /* this statement adds mPAW to each case */
  IF _n_ = 1 THEN SET forwgtmean ;
  SET libname.skinnydataset ;
  /* normalize each individual's weight */
  NPAW = PAW/mPAW;
RUN;
```

2. Mixed (subject-specific) multi-level models. In these models, the aim is to not merely control for within-person correlation, but to purposively

model both between- and within-person variance. For such models, it is recommended that analysts create a baseline survey weight (base weight) at the person level using each respondent's earliest panel analysis weight, followed by calculating adjusted wave-level weights, and then assigning the relevant weights as appropriate in multi-level analysis syntax.

For further details, please see Ch. 11 in Heeringa, West, and Berglund (2017).[10] The final person-level base weights should be normalized prior to analysis such that they sum to the *n* of respondents (not waves) in the analytic sample; the final wave-level weights should be normalized such that they sum to the *n* of waves within respondent.

Recommended Citation

Patrick, M. E., Terry-McElrath, Y. M., Berglund, P., Pang, Y. C., Heeringa, S. G., & Si, Y. (2022). *An Updated Weighting Strategy for the Monitoring the Future Panel Study.* Monitoring the Future Occasional Paper No. 98. University of Michigan Institute for Social Research: Ann Arbor, MI. doi:10.7826/ISR-UM.06.585140.002.08.0001.2022

MonitoringtheFuture.org **@ysi_news** **MTFinformation@umich.edu**

[10] Heeringa, S. G., West, B. T., & Berglund, P. A. (2017). *Applied Survey Data Analysis* (2nd ed.). CRC Press.